

UNIVERSIDAD NACIONAL DE PIURA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE ESTADISTICA



**“AJUSTE DE LA CURVA LOGISTICA APLICADA EN LOS
DATOS CENSALES DE LA POBLACION PERUANA, 1836-2007”**

Presentada por:

Brisset Melissa Cruz Ipanaqué
Corona del Rosario Rosas Torres

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE
LICENCIADA EN ESTADISTICA.**

Piura, Perú

Marzo 2015

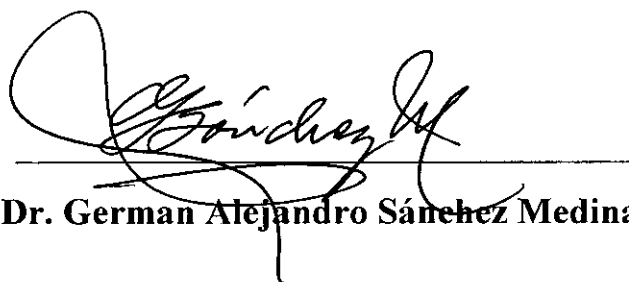
7620
CRU



**TESIS PRESENTADA COMO REQUISITO PARA OPTAR
EL TITULO PROFESIONAL DE:**

LICENCIADA EN ESTADISTICA

Presidente de Jurado de Tesis:



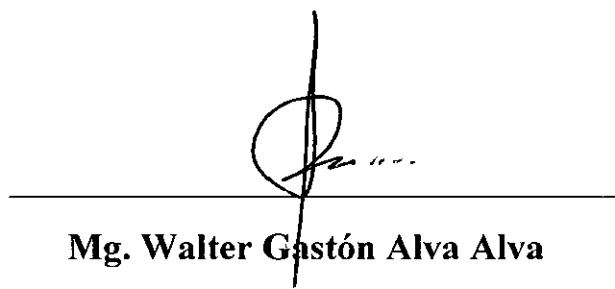
Dr. German Alejandro Sánchez Medina

Secretario de Jurado de Tesis:



Mg. Lemín Abanto Cerna

Vocal de Jurado de Tesis:



Mg. Walter Gastón Alva Alva

DEDICATORIAS

Esta tesis se la dedico a mi Dios quien supo guiarme por el buen camino, darme la fuerza para seguir adelante y no desmayar en los problemas que se presentaban, enseñándome a encarar las adversidades sin perder nunca la dignidad ni desfallecer en el intento

A mis padres, por su amor, trabajo y sacrificio en todos estos años, gracias a ustedes he logrado llegar a esta instancia de mis estudios y convertirme en lo que soy. Con su amor y enseñanza han sembrado las virtudes que se necesitan para vivir con anhelo y felicidad. Ha sido un privilegio ser su hija, son los mejores padres.

En un sincero e intenso amor, para la persona que con su apoyo constante y amor incondicional ha sido mi amigo y compañero inseparable. Gracias Gerardo, porque sin proponértelo le entregaste a mi corazón el mensaje más precioso de amor.

A mi preciosa hija Brihanna para quien ningún sacrificio es suficiente, que con su luz ha iluminado mi vida y hace mi camino más claro.

Brisset M. Cruz Ipanaqué

A mi Dios por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A mis ángeles que en todo momento al desarrollar esta investigación estuvieron conmigo, por eso a ti Talía hermana querida te dedico mi esfuerzo donde te encuentres.

A mis padres, por su apoyo incondicional, motivación y comprensión, en todo momento y siempre estuvieron brindándome su ayuda. A mi hermana Julissa por sus palabras de aliento. A mi sobrina Gaela, para que vea en mi un ejemplo a seguir.

A mis amigas, a quienes fueron no solo mis amigas sino mis hermanas, gracias por reír y llorar conmigo, me supieron escuchar y dar su apoyo. Y en especial a Emily Braekhus por su apoyo incondicional.

A Brisset, mi compañera y amiga de tesis, quien es un ejemplo a seguir por la dedicación y compromiso para concluir con este logro.

Corona del R. Rosas Torres

AGRADECIMIENTOS

En la presente investigación nos gustaría agradecerle a ti Dios por bendecirnos y permitirnos hacer realidad esta meta trazada.

A nuestro Asesor de tesis, Dr. Conrado Vargas Lynch por su esfuerzo y dedicación, quien con su arduo trabajo de transmitirnos sus diversos conocimientos, su experiencia, su paciencia y su motivación logro convertirnos en profesionales de éxito.

Igualmente agradecer a mis profesores durante toda mi carrera profesional porque todos han aportado con un granito de arena a nuestra formación, y en especial al Dr. German Sánchez Medina y Lic. Lemin Abanto Cerna por sus consejos, su enseñanza y más que todo por su amistad.

A nuestros padres fuente de apoyo constante e incondicional en toda nuestra vida y más aún en los duros años de carrera profesional, quienes nos han apoyado y creído en nosotras, sabiéndonos encaminar con sus sabios consejos para lograr nuestras metas. A nuestros hermanos por motivarnos. A ustedes por siempre nuestro corazón y respeto.

Son muchas las personas que han formado parte de nuestra vida profesional que nos encantaría agradecerles su amistad, consejos, apoyo, ánimo y compañía en los momentos más difíciles de nuestra vida.

Algunas están aquí y otras en nuestros recuerdos y corazón, sin importar en donde estén darles las gracias por formar parte de nosotras, y por todo lo que nos han brindado.

Muchas gracias y que Dios los bendiga.

Las autoras

RESUMEN

Según el informe del Fondo Alemán de la Población de la Tierra (2014), la población sobrepasa los siete mil doscientos millones de personas, lo que nos acarrea a plantearnos la siguiente pregunta, ¿Cuántas personas puede apoyar la tierra y en qué condiciones?, convirtiéndose esto en una situación alarmante. Uno de los principales conflictos que tiene que enfrentar un país es el crecimiento poblacional; la idea de que este incremento es imparable, el miedo a la insuficiencia de recursos, el hambre, son problemas que preocupan cada vez más al hombre de hoy y le hacen preguntarse por su futuro.

El objetivo de esta investigación es: Ajustar la función logística a los datos censales de la población peruana en un periodo de tiempo comprendido entre los años 1836 – 2007.que permita obtener parámetros de referencia para evaluar la cobertura de eventos censales futuros; de ahí la importancia de encontrar un modelo que proporcione algún criterio de evaluación, un modelo que describa el cambio en el tamaño de la población a través del tiempo y que sea lo más confiable

El resultado de la investigación muestra la metodología empleada, los ajustes necesarios, la tasa relativa de crecimiento poblacional, estimaciones puntuales para un periodo determinado de años y el coeficiente de correlación, finalmente se logra obtener un modelo con un buen ajuste estadístico.

Palabras claves: Modelos matemáticos de población, Demográfica, censos, Crecimiento logístico.

ABSTRACT

According to the report of the German Foundation for World Population (2014), the population exceeds seven thousand two hundred million people, which raises the question of how many people the Earth can support and under what conditions, making it an alarming situation. One of the main conflicts that a country has to deal is the population growth; the idea that this increase is unstoppable, the fear about lack of resources and hunger are current growing problems that make us wonder about our future.

The objective of this research was to adjust the logistic function to the census data of Peruvian population over a period of time from 1836 to 2007 that allows to obtain reference parameters for assess the coverage of future census events. It is very important to find a reliable model that provides some evaluation criteria and can also describes the change in population size over time.

The result of the research shows the methodology used, the necessary adjustments, the relative rate of population growth, point estimates for a given period of years and the correlation coefficient. Finally a model with a good statistical adjustment was obtained.

Keywords: Mathematical models of demographic population, censuses, logistic growth.

INTRODUCCION

El interés del país exige que su población no crezca desmesuradamente y que no disminuya en forma alarmante. Hay dos razones: un país despoblado es un país muerto y la sobrepoblación es un factor que acarrea e incrementa conflictos sociales. Con el fin de describir las diferentes posiciones de los gobiernos respecto al crecimiento de la población, a lo largo de la historia de la ciencia demográfica se han desarrollado diversas funciones aplicadas a la población total.

En la presente investigación se construyó una función matemática con el propósito de analizar la dinámica demográfica de Perú. No se trata de aplicar o comparar modelos ya existentes, sino de generar uno nuevo a partir de una función matemática de tipo logístico que muestra la dinámica de la tasa de crecimiento demográfico, resolviendo para ello la ecuación diferencial que se produce.

La finalidad es ajustar la función logística a los datos censales de la población peruana en un periodo de tiempo comprendido entre los años 1836 – 2007, que permita obtener parámetros de referencia para evaluar la cobertura de eventos censales.

La gran utilidad de estos modelos han conformado un instrumento de vital importancia para todas las políticas y programas de los gobiernos centrales y locales, a la vez que permiten anticipar las demandas sociales básicas relativas en educación, empleo, vivienda, salud y seguridad social, y así encaminar los criterios de asignación de recursos.

Las autoras

INDICE

DEDICATORIA -----	4
AGRADECIMIENTO-----	5
RESUMEN-----	6
INTRODUCCION-----	7
I. PLANTEAMIENTO DE PROBLEMA -----	11
1 PLANTEAMIENTO DE PROBLEMA-----	11
1.1.1 REALIDAD PROBLEMÁTICA -----	11
1.1.2 FORMULACIÓN DE PROBLEMA -----	13
1.2 JUSTIFICACIÓN-----	14
1.3 OBJETIVO -----	14
1.3.1 OBJETIVO GENERAL-----	15
1.3.2 OBJETIVO ESPECÍFICO -----	15
1.4 PRUEBA DE HIPOTESIS-----	15
II. MARCO TEÓRICO-----	16
2.1 ANTECEDENTES-----	16
2.1.1 ANTECEDENTES INTERNACIONALES-----	16
2.2 CURVA LOGISTICA -----	20
2.2.1 ANTECEDENTES-----	20
2.2.2 FUNCIÓN LOGISTICA: DERIVACIÓN DEL MODELO MATEMATICO -----	21
2.2.3 RÉGIMEN DE CRECIMIENTO LOGÍSTICO-----	25
III. METODOLOGÍA -----	37
3.1 TIPO DE INVESTIGACIÓN-----	37
3.2 POBLACIÓN Y MUESTRA -----	38

IV. RESULTADOS DE LA INVESTIGACIÓN	38
4.1 ESTIMACIÓN DE PARÁMETROS	38
V. CONCLUSIONES	59
5.1 CONCLUSIONES	59
5.2 RECOMENDACIONES	60
VI. BIBLIOGRAFIA	61
ANEXO 1	62
ANEXO 2	65
ANEXO 3	66

I. PROBLEMA DE INVESTIGACION

1.1. PLANTEAMIENTO DEL PROBLEMA

1.1.1. Realidad problemática

Según el informe del Fondo Alemán de la Población de la Tierra (2014), la población sobrepasa los siete mil doscientos millones de personas, lo que nos acarrea a plantearnos la siguiente pregunta, ¿Cuántas personas puede apoyar la tierra y en qué condiciones?, convirtiéndose esto en una situación alarmante. Uno de los principales conflictos que tiene que enfrentar un país es el crecimiento poblacional; la idea de que este incremento es imparable, el temor a la progresiva destrucción del planeta, el miedo a la insuficiencia de recursos, el hambre, son problemas que preocupan cada vez más al hombre de hoy y le hacen preguntarse por su futuro.

En América latina la tasa de crecimiento demográfico se ha incrementado rápidamente, aunque no de manera uniforme ni predecible. Ello puede atribuirse a la baja continua de la mortalidad y a persistentes tasas de elevada fecundidad. La combinación de los avances en la producción de alimentos, la tecnología médica y la infraestructura de saneamiento han resultado en nacimientos humanos que superan por mucho las muertes en un nivel sin precedentes en la historia humana.

El interés del país exige que su población no crezca desmesuradamente y que no disminuya en forma alarmante. Hay dos razones: un país despoblado es un país muerto y la sobrepoblación es un factor que acarrea conflictos sociales. Con el fin de describir las diferentes posiciones de los gobiernos respecto al crecimiento de la población, a lo largo de la historia de la ciencia demográfica se han desarrollado diversas funciones, como la logística, aplicada a la población

total. Los esfuerzos por establecer leyes matemáticas que representen el crecimiento demográfico se han acrecentado gracias a la creciente disponibilidad de estadísticas y al desarrollo de modelos matemáticos para la descripción de datos.

Los métodos matemáticos que se aplican en el cálculo de la población futura del país, se basan en ecuaciones que expresan el crecimiento demográfico en función del tiempo, dicho crecimiento medido y expresado en una tasa o en un porcentaje de cambio, se obtiene a partir de la observación o estimación del volumen poblacional en dos o más fechas del pasado reciente. Por lo general, los censos de población, realizados con un intervalo aproximado de diez años, permiten dicha medición.

Uno de los modelos matemáticos empleados en la demografía para analizar las tendencias del crecimiento de una población y de diversos indicadores demográficos es la función logística. El ajuste de la curva logística a una población observada, se fundamenta en el principio de que: la población atraviesa primero, por una fase de desarrollo más bien lenta, luego crece rápidamente, después se detiene más y más hasta que la población se vuelve estacionaria. Por consiguiente, si se acepta la premisa de que las poblaciones no tienen un continuo crecimiento indefinido a través de períodos extensos, ni la curva polinomial, ni la geométrica, expresarían fielmente las variaciones y alternativas, particularmente si el análisis cronológico se realiza con el fin de pronosticar la población a largo plazo.

Por otro lado, actualmente existe la necesidad de obtener parámetros de referencia que permitan evaluar la cobertura de eventos censales futuros; de ahí la importancia de encontrar un modelo que proporcione algún criterio de evaluación, un

modelo que describa el cambio en el tamaño de la población a través del tiempo y que sea lo más confiable posible.

Pero una de las dificultades más importantes para estimar la población futura, es que muchos casos solo es posible obtener datos incompletos y a veces imprecisos de la evolución demográfica y el desarrollo económico de la localidad. La imposibilidad de llevar a cabo continuos levantamientos censales para suplir estas necesidades de información, debido a los enormes esfuerzos que ello implica tanto técnicos como logísticos, y obviamente financieros, conlleva a recurrir al diseño de modelos que permitan disponer de información sobre los probables volúmenes y diversas características de la población futura. La gran utilidad de estos modelos han conformado un instrumento de vital importancia para todas las políticas y programas de los gobiernos centrales y locales, a la vez que permiten anticipar las demandas sociales básicas relativas en educación, empleo, vivienda, salud y seguridad social, y así encaminar los criterios de asignación de recursos.

1.1.2. Formulación del problema

En este trabajo se propone dar respuesta a la Cuestión:

¿Es posible estimar una función logística a los datos censales de la población peruana en un periodo de tiempo comprendido entre los años 1836 - 2007?

1.2. JUSTIFICACION

En tal sentido, la presente investigación se justifica a partir de los siguientes puntos de vista:

En el ámbito teórico, existen pocos estudios e investigaciones que aborden en forma integral un modelo matemático del crecimiento poblacional en el Perú. La falta de un estudio actualizado y pertinente, amerita un trabajo de investigación que centre su esfuerzo en mostrar que el cambio en el tamaño de la población peruana, en el periodo 1836 - 2007, tiene un comportamiento como el que establece la ley logística de crecimiento poblacional.

En el aspecto práctico, la curva logística contribuye al buen ajuste de los datos del crecimiento de la población peruana y provee un modelo aceptable para explicar su dinámica poblacional, que permiten anticipar las demandas sociales básicas relativas en educación, empleo, vivienda, salud y seguridad social, y así encaminar los criterios de asignación de recursos.

En el aspecto metodológico, es importante porque es un modelo que permite describir y pronosticar el crecimiento de la población peruana, a lo largo del tiempo, además constituye un elemento de valor fundamental no solo para los usuarios privados, sino también para los sectores gubernamentales, y en particular para las áreas de planificación y desarrollo, ya que permiten conocer la probable evolución cuantitativa de la población para propiciar y orientar la transformación cualitativa y el desarrollo integral de la sociedad futura.

1.3. OBJETIVOS

1.3.1. Objetivo general

- Ajustar la función logística a los datos censales de la población peruana en un periodo de tiempo comprendido entre los años 1836 – 2007.

1.3.2. Objetivos específicos

- Estimar los parámetros de la función logística L, K, b .
- Estimar el Coeficiente de correlación.
- Realizar el pronóstico puntual para la población de los años 2015-2060.

1.4. HIPOTESIS DE INVESTIGACION

Si es posible estimar una función logística a los datos censales de la población peruana en un periodo de tiempo comprendido entre los años 1836 - 2007.

II. MARCO TEORICO

2.1. ANTECEDENTES GENERALES

2.1.1. Antecedes internacionales

Ordorica (2008), construyó una función matemática capaz de describir la dinámica de la tasa de crecimiento demográfico de la población de México y, en consecuencia, la evolución de la población entre 1980 y 2005, y sus perspectivas, para establecer escenarios de lo que podría ocurrir en el corto, mediano y largo plazos.

Al evaluar sobre la función matemática obtuvo las siguientes conclusiones:

- Las tasas de crecimiento demográfico para México estimadas por el Conapo, de .025 para 1980, de .022 para 1985, de .019 para 1990, de .016 para 1995, de .013 para 2000 y de .009 para 2005. La población inicial calculada por Conapo para 1980 fue de 67.4 millones. El coeficiente de correlación del modelo fue de .997 y el de determinación de .994, lo que muestra un excelente ajuste.
- Las proyecciones de población son válidas en periodos cortos debido a la inercia demográfica, y que estimaciones a largo o muy largo plazos deben considerarse especulaciones que sólo nos sirven para analizar los diversos escenarios que podrían presentarse.
- El modelo desarrollado proporciona resultados adecuados en el ámbito municipal tratándose de áreas pequeñas y, sobre todo, es más fácil establecer hipótesis sobre las cifras futuras en el crecimiento demográfico

Poveda & Manrique (2007), realizaron una investigación sobre aplicación de la curva logística a los censos de la ciudad de Medellín - Colombia. Donde concluye que:

- La curva logística permite pronosticar valores de población hacia el futuro y también, hacia el pasado, pues, con la ecuación logística solo deben reemplazarse los valores numéricos del año en el que se desea hacer la estimación demográfica.
- El método de cálculo utilizado es sencillo y los valores calculados presentan buen ajuste con los datos reales por lo tanto, con este método, la curva logística se ajusta muy bien a los datos de los censos realizados entre los años 1912 – 2005 para la ciudad de Medellín – Colombia.
- El coeficiente de correlación o de conformidad $r^2=0,99999997$ es altamente satisfactorio, lo cual confirma la precisión del método.
- Hacia el año 2030, la población de la ciudad de Medellín (sin sus municipios vecinos de hoy 2007) habrá alcanzado los tres millones de personas, según los datos existentes a la fecha.
- La población seguirá aumentando durante los próximos años, aun cuando con tasas porcentuales declinantes

Medina & Suárez (2000), realizaron una investigación sobre el Modelo logístico adaptado a la población mexicana en el periodo 1900 – 1990. Para el desarrollo de este trabajo utilizamos el método de “Variables separables” de las ecuaciones diferenciales ordinarias de primer orden, así como una hoja de cálculo para obtener los valores de los parámetros del “modelo logístico” recurriendo al método de mínimos cuadrados. Obteniendo como valor de $K = a/b = 200.210526$ el cual representa la población límite, es decir la población máxima esperada a largo plazo. Y los valores de los parámetros fueron $a = 0.03804$ y $b = 0.00019$ que corresponden a los datos poblacionales registrados en los años 1950, 1970 y 1990 para estos valores se obtuvo un coeficiente de correlación de 0.99278, es decir, más de 99.2% de los datos observados son explicados por el modelo.

Se arribó a las siguientes conclusiones:

- Este método ofrece la ventaja de que con tan solo una terna de datos se pueden obtener los parámetros del modelo. Además de que entre más datos observados se incorporen al análisis mayor precisión se obtiene en la aproximación.
- De acuerdo a los datos obtenidos podemos afirmar que la evolución de la población mexicana tiene un comportamiento que va de acuerdo con la ley logística de crecimiento de la población pero restringida al periodo 1900-1990.
- Previamente a su aplicación en otras poblaciones, se considera importante realizar un estudio a mayor detalle donde se vincule con indicadores demográficos que lo justifique y donde se efectúen algunas comparaciones con otros modelos.

Veres (1985), realizó un estudio sobre el Nuevo procedimiento para el ajuste de la curva logística: aplicación a la población española. Donde se presenta un procedimiento de ajuste de la curva logística, basado en el conocimiento previo del punto de inflexión. Dicho procedimiento proporciona como estimación las asíntotas superior e inferior de la curva. Para realizar la investigación se basan en un estudio previo de las tasas de crecimiento intercensales experimentadas por la población española a lo largo de la historia de sus censos de población desde 1857 hasta 1981. Al evaluar se pudo concluir que:

- La logística proporciona buenas previsiones sobre la cuantía futura de las poblaciones humanas, siendo indudable su importancia histórica en el campo de la demografía.
- La principal ventaja del procedimiento de ajuste realizado en este estudio, aparte de su rapidez y operatividad, reside en su adaptabilidad a las posibles modificaciones que se producen con la introducción de nuevos datos censales. En efecto, cada nuevo resultado censal puede confirmar o no el ajuste realizado para la obtención del punto de inflexión, dando lugar a una posible modificación de la logística para adaptarla mejor a la nueva situación.
- El análisis previo del ciclo en el que se encuentra la población es fundamental para proceder a ligeras modificaciones de la logística.
- El procedimiento de ajuste desarrollado evita la estimación a priori de las asíntotas, sustituyéndola por la del punto de inflexión, siendo la naturaleza de los datos disponibles la que determine la conveniencia o no del empleo del ajuste que se ha considerado.

2.2. CURVA LOGISTICA

2.2.1. Antecedentes

La curva logística es suficientemente conocida entre estadísticos, biólogos y demógrafos como modelo determinístico para describir y para pronosticar el crecimiento de una población humana, a lo largo del tiempo, cuando están sujetas a ciertas condiciones de su régimen de crecimiento.

En propias palabras de Robert Malthus en su Ensayo sobre el principio de la población 1798: La población, si no encuentra obstáculos, aumenta en progresión geométrica, mientras que los alimentos lo hacen en progresión aritmética. Aunque las tesis de Malthus estuvieron alejadas de la realidad, tuvieron una gran influencia en la política. El pronosticaba que a partir de cierto límite de la población habría escasez de alimentos y en consecuencia una gran hambruna. Dicha hambruna jamás se produjo debido al desarrollo tecnológico, por lo que los supuestos de su modelo resultaron simplistas. Una de las críticas más fuertes al modelo de Malthus radica en suponer que los recursos del medio son inagotables. Con un sentido más realista, puede decirse que efectivamente los recursos están limitados y que las poblaciones de ninguna manera pueden crecer indefinidamente al ritmo malthusiano.

El belga Adolphe Quetelet fue uno de los primeros en utilizar un modelo matemático (1835) para explicar la evolución de la población, señalando que la población se incrementa a ritmo acelerado hasta llegar a un punto en que empieza a crecer más lentamente.

Como influencia de las teorías de Malthus en lo referente al desequilibrio entre el incremento de la población y el de los alimentos, Darwin supuso (1838) la necesidad de los seres vivos de competir para obtener suficiente alimento que les garantice la mejor situación posible en el entorno.

Verhulst analizo este planteamiento y sugirió una curva teórica a la que llamo logística, la cual dio a conocer en 1837, y actualmente es conocida como Ley logística del crecimiento de la población.

Esta curva es un refinamiento del crecimiento exponencial y adopta la forma de una S invertida. Cuando una magnitud crece en un sistema finito, a partir de cierto momento el tamaño del sistema limita el crecimiento de la magnitud al no haber recursos abundantes suficientes para seguir permitiendo el crecimiento exponencial. En un comienzo Verhulst supuso que los obstáculos aumentan “exactamente en la misma proporción que la población superabundante”, pero más tarde reemplazó este supuesto con la hipótesis de que los obstáculos “aumentan en proporción al porcentaje del exceso de población respecto de la población total” (Verhulst, 1938 y 1945).

Usando esta ley encontró dos modelos, con los cuales pronostico para Bélgica una población máxima de 6 millones y para Francia de 40 millones. En 1930, la población de Francia concordaba sorprendentemente con lo obtenido por Verhulst.

La función logística permaneció olvidada por muchos años debido a la falta de datos. En 1920, sin conocer los trabajos de aquellos científicos europeos, dos biometristas americanos, R. Pearl & L. Reed redescubrieron la función logística y la utilizaron para explicar el comportamiento del crecimiento de la población en Estados Unidos.

2.2.2. Función logística: derivación del modelo matemático.

La densidad poblacional de una especie única en el tiempo t será denotada por $x(t)$, y $\frac{dx}{dt}$ o $x'(t)$ la tasa de cambio de tamaño de la población. Asumiremos que la tasa de crecimiento de la población depende solamente del tamaño de la población. Tal suposición parece ser razonable para los organismos simples como los microorganismos. Para los organismos más complejos, como las

plantas, los animales o los seres humanos esto es obviamente una simplificación excesiva ya que ignora la competencia entre especies por los recursos, así como otros factores importantes, como la estructura de edad (la tasa de mortalidad puede depender de la edad y no en la densidad de población, mientras que la tasa de natalidad puede depender del tamaño de la población adulta en lugar de tamaño de la población total)

Además, la posibilidad de que las tasas de nacimiento o defunción pueden ser influenciados por el tamaño de las poblaciones que interactúan con la población objeto de estudio también debe ser considerado (competencia, depredación, mutualismo).

La tasa de crecimiento per cápita, o la tasa de crecimiento por cada miembro, está dada por $x'(t)/x(t)$, que estamos asumiendo es una función de $x(t)$.

Se considera que las tasas de crecimiento total disminuyen a medida que aumenta el tamaño de la población

El modelo de población más simple en el que la tasa de crecimiento per cápita es una función decreciente del tamaño de la población es $\lambda - ax$. Esta suposición conduce a la ecuación diferencial logística

$$x' = x(\lambda - ax)$$

Introducido por primera vez por Verhulst (1838) y además más tarde estudiado por R. Pearl & L.J Reed (1920). Esta ecuación es escrita comúnmente en la forma

$$x' = rx \left(1 - \frac{x}{K}\right) \quad (1)$$

Con los parámetros $r = \lambda$, $K = \lambda/a$. Los parámetros r y K , asumido positivo, puede dar entonces importancia biológica. Se observa que $x' \approx rx$ cuando x es pequeño, y que $x' = 0$ cuando x está cerca de K . En otras palabras, cuando es pequeña la población experimenta un crecimiento exponencial, mientras que cuando x está cerca de K la población apenas cambia.

Separación de variables nos permite reescribir la ecuación

(1) como:

$$\int \frac{dx}{x(K-x)} = \frac{r}{K} \int dt$$

Usando fracciones parciales, da

$$\frac{1}{x(K-x)} = \frac{1}{K} \left(\frac{1}{x} + \frac{1}{K-x} \right)$$

Nos permite integrarlo:

$$\begin{aligned} \frac{r}{K} t + c &= \int \frac{dx}{x(K-x)} = \frac{1}{K} \left(\int \frac{dx}{x} + \int \frac{dx}{K-x} \right) \\ &= \frac{1}{K} (\log x - \log(K-x)) \end{aligned}$$

Donde c es la constante de integración.

Si el tamaño de la población en el momento $t = 0$ es x_0 la sustitución de la condición inicial $x(0) = x_0$ da

$$c = \frac{1}{K} (\log x_0 - \log(K-x_0))$$

Ahora tenemos

$$\begin{aligned} \frac{1}{K} (\log x - \log(K-x)) &= \frac{r}{K} t + \frac{1}{K} (\log x_0 - \log(K-x_0)) \\ &= \frac{r}{K} t + \frac{1}{K} (\log x_0 - \log(K-x_0)) \end{aligned}$$

$$\log \left(\frac{x}{K-x} \right) = rt + \log \frac{x_0}{K-x_0}$$

$$\log \frac{x(K-x_0)}{x_0(K-x)} = rt$$

$$\frac{x(K-x_0)}{x_0(K-x)} = e^{rt}$$

Además la simplificación algebraica da

$$x(K-x_0) = x_0(K-x)e^{rt} = Kx_0e^{rt} - xx_0e^{rt}$$

$$x(K-x_0 + x_0e^{rt}) = Kx_0e^{rt}$$

Y finalmente

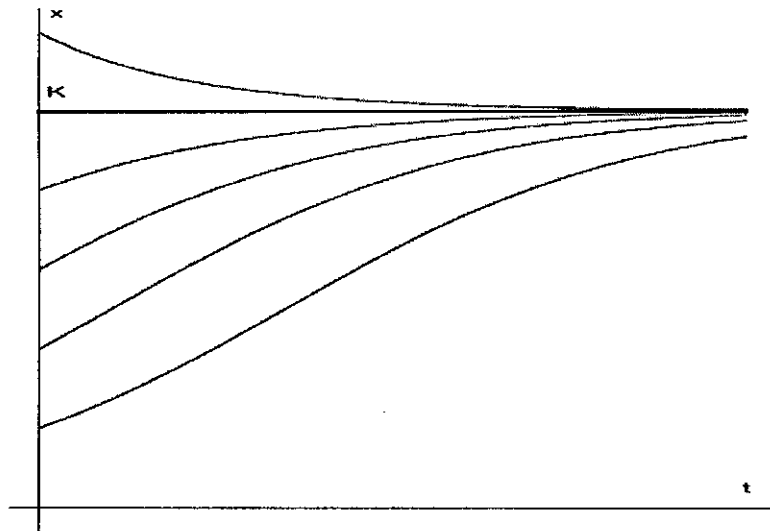
$$x(t) = \frac{Kx_0e^{rt}}{K-x_0+x_0e^{rt}} = \frac{Kx_0}{x_0+(K-x_0)e^{-rt}} \quad (2)$$

La solución anterior es válida sólo si $0 < x_0 < K$, de manera que los logaritmos obtenidos en la integración se definen. Para obtener la solución sin esta restricción, nuestra integración debería haber dado los logaritmos de los valores absolutos. Sin embargo, la ecuación (2) para la solución de la ecuación logística es válida para todos x_0 , como podría ser verificado por un análisis más cuidadoso.

La ecuación (2) para la solución del problema logístico valor inicial muestra que el tamaño de la población $x(t)$ se acerca al límite K cuando $t \rightarrow \infty$ si $x_0 > 0$. El valor de K se llama la capacidad de carga de la población, porque representa el tamaño de la población de los recursos disponibles pueden seguir apoyando. El valor r se denomina tasa intrínseca de crecimiento, ya que representa la tasa de crecimiento del por lograr para que el tamaño de la población eran lo suficientemente pequeño para asegurar la limitación de recursos insignificantes. El modelo logístico predice un rápido crecimiento inicial para $0 < x_0 < K$, a continuación, una disminución en la tasa de crecimiento como el tiempo pasa para que el tamaño de la población se aproxime a un límite (Grafica 1). Este comportamiento está de acuerdo con el comportamiento observado de muchas poblaciones, y por esta razón, el modelo logístico se utiliza a menudo como un medio de describir la población tamaño.

Grafica 1

Solución de la ecuación logística



2.2.3. Régimen de crecimiento logístico

Según Poveda & Manrique (2007), Para un mejor entendimiento de la aplicación, se transcribe la teoría que soporta el método.

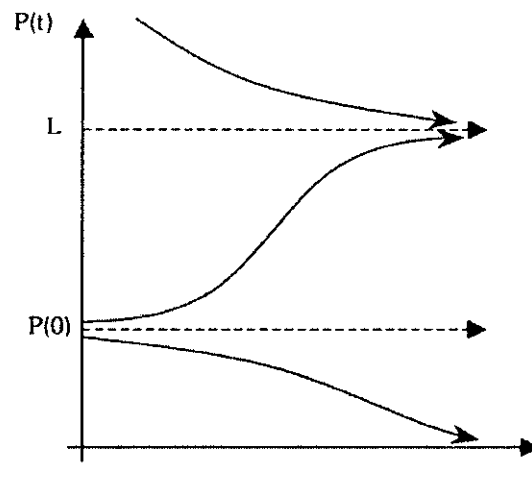
1. El régimen de crecimiento logístico resulta cuando se supone que la dinámica de crecimiento en el tiempo está determinada por tres hipótesis básicas, a saber:
 - a. El crecimiento vegetativo (más las inmigraciones, si las hay), es en todo momento proporcional al número de individuos presentes, al cual llamaremos p .
 - b. La población vive en un espacio limitado y/o bajo condiciones externas restrictivas que no le permitirán, ni aun a largo plazo, exceder de un tamaño límite, al cual llamaremos L .

- c. En todo momento la velocidad de crecimiento es proporcional al factor $(L - p)$ que representa las fuerzas exógenas restrictivas.

En la Gráfica 2, se observan los comportamientos creciente y decrecientes de la población en el tiempo según las condiciones iniciales de: superpoblación desde $P(t)$ hacia una población límite L ; de crecimiento ascendente desde una población inicial $P(0)$ hasta la población límite L ; y de crecimiento negativo o decrecimiento, desde una población crítica hasta la extinción.

Gráfica 2.

Comportamientos de la Curva logística



Dentro de las hipótesis anteriores, y si se sabe que el dato censal que es históricamente más reciente corresponde al valor más alto ya censado, se determina que la ecuación diferencial para el crecimiento para una población ascendentemente creciente, es:

$$\frac{dp}{dt} = bp(L - p) \quad (3)$$

En donde:

p : Número de individuos de la población que viven en el instante t del tiempo.

b : Coeficiente constante y positivo que es propio de las características vitales de la población y del régimen biológico y ecológico en que vive.

L : Población límite posible, que está determinada por el entorno vital de la población.

Integrando la ecuación diferencial (3) ya escrita, se obtiene la función de crecimiento de la población:

$$\hat{P}(t) = \frac{L}{1 + Ke^{-bt}} \quad (4)$$

En donde K es un constante que se introduce al efectuar la integración y que, numéricamente, depende del volumen que tenga la población en el momento que se tome como instante inicial ($t = 0$) para comenzar a contar el tiempo. En efecto, de la ecuación (4) se deduce que

$$K = \frac{L}{p(0)} - 1$$

Es evidente que la forma de la ecuación (4) es invariante ante un cambio en el origen de tiempo y ante un cambio en la unidad de medida del tiempo. En efecto, haciendo un cambio en ambos sentidos, es decir, tomando una nueva variable temporal

$$T = ct + a,$$

Donde a expresa un cambio en el origen del tiempo (o cronología) y c un cambio en la unidad de medida del tiempo (o cronometría), la ecuación toma la forma:

$$\hat{P}(t) = \frac{L}{1 + K_1 e^{-Bt}}$$

En donde:

$$K_1 = Ke^{ab/c} \text{ y } B = b/c$$

2. En las aplicaciones demográficas de la logística, se trata usualmente de “ajustar” una función de este tipo a una sucesión consecutiva de valores numéricos de la población p_1, p_2, \dots, p_n ,

Los cuales son medidos experimentalmente (por ejemplo, mediante recuentos o censos de la población, en momentos sucesivos y distintos del tiempo

$$t_1, t_2, \dots, t_n$$

Los cuales se disponen en orden creciente, pero no son necesariamente equidistantes.

En tales situaciones los datos de población (p_1, p_2, \dots, p_n) suelen ser crecientes en el tiempo. Estos datos presentan tres características prácticas, que son importantes para entender el problema teórico del ajuste de la curva logística:

- a. Casi inevitablemente están afectados por errores de recuento o pueden ser el resultado de estimaciones y no de enumeraciones precisas.
- b. Aunque tales datos fueran rigurosamente exactos, no puede esperarse que todos caigan rigurosa y exactamente sobre una función basada en hipótesis determinísticas ya que en el crecimiento de un agregado colectivo, como en el caso de una población biológica, hay inevitablemente desviaciones aleatorias más o menos apreciables con respecto a la función determinística.
- c. Sucede entonces que, aun cuando la realidad geográfica, ambiental, ecológica y social justifique la adopción de una

curva logística de crecimiento, el problema del ajuste a los datos empíricamente observados en el tiempo, presenta varias dificultades de técnica numérica. Ese problema consiste, en esencia, en calcular los tres parámetros L, k, b de la ecuación (4), a partir de la serie de recuentos, que suele darse en forma tabulada.

Es evidente que, para hacer el ajuste, no se puede esperar que la curva -la cual sólo tiene tres parámetros para estimar- pase exactamente por los puntos tabulados, lo cual daría lugar a n condiciones y como, en general, n es distinto de 3, el problema no tendría solución. Aun el caso de tener tres observaciones censales ($n = 3$), las observaciones teórico-prácticas mencionadas descartan este método.

El ajuste de la logística por mínimos cuadrados

3. La ecuación (4) de la curva logística puede escribirse en la forma equivalente:

$$\ln\left(\frac{L}{p} - 1\right) = \ln(K) - bt$$

O bien como

$$z = A - bt, \quad (5)$$

Siendo

$$z = \ln\left(\frac{L}{p} - 1\right) \quad (6)$$

$$A = \ln(k) \quad (7)$$

Y donde todos los logaritmos son neperianos (naturales).

Lo anterior significa que si, en lugar de representar la ley logística en coordenadas cartesianas e isométricas para t en las abscisas y p

en las ordenadas, la representamos en coordenadas cartesianas o isométricas para t en abscisas y z en ordenas; si a L se le atribuye el valor correcto, se obtendrá una línea recta, que representa a la ecuación (5). Si los valores de z se calculan en la ecuación (6) con valores de L que no son los verdaderos valores de la ecuación (5), la gráfica entre t y z no dará una recta, sino una curva.

Esto mismo equivale a decir que, calculando z con el valor L correcto de la logística, entonces la gráfica de z versus t es una parábola

$$z = A + Bt + Ct^2 \quad (8)$$

En donde $C = 0$, siendo $A = \ln(K)$ y $B = -b$

Ahora bien, es sabido en estadística elemental, que la parábola

$$z = A + Bt + Ct^2$$

Puede ajustarse a una serie $\{t_i, z_i\}$ de datos empíricos, usando el criterio de mínimos cuadrados, y su método de análisis. En ese caso, las ecuaciones normales que determinan los tres parámetros A, B, C son las del sistema

$$\begin{aligned} \sum z_i &= An + B \sum t_i + C \sum t_i^2 \\ \sum z_i t_i &= A \sum t_i + B \sum t_i^2 + C \sum t_i^3 \\ \sum z_i t_i^2 &= A \sum t_i^2 + B \sum t_i^3 + C \sum t_i^4 \end{aligned}$$

En donde n es el número de fechas censales que se tienen. En todos los casos de aplicación práctica se comprueba que el sistema es resoluble, o sea que el sistema no es homogéneo (es decir, los tres términos

$\sum z_i, \sum z_i t_i, \sum z_i t_i^2$ no se anulan simultáneamente), y que el determinante del sistema

$$D \equiv \begin{vmatrix} n & \sum t_i & \sum t_i^2 \\ \sum t_i & \sum t_i^2 & \sum t_i^3 \\ \sum t_i^2 & \sum t_i^3 & \sum t_i^4 \end{vmatrix}$$

Es distinto de cero.

Para simplificar la escritura, pondremos:

$$T_1 = \sum t_i, \quad T_2 = \sum t_i^2, \quad T_3 = \sum t_i^3, \quad T_4 = \sum t_i^4$$

Aplicando a las ecuaciones normales la regla de Cramer, se obtiene para el coeficiente C la expresión

$$C = \frac{\begin{vmatrix} n & T_1 & \sum z_i \\ T_1 & T_2 & \sum z_i t_i \\ T_2 & T_3 & \sum z_i t_i^2 \end{vmatrix}}{D}$$

Para la curva logística, con L correcto, se debe tener $C = 0$, como ya se dijo. Por tanto, se debe tener la condición

$$\begin{vmatrix} n & T_1 & \sum z_i \\ T_1 & T_2 & \sum z_i t_i \\ T_2 & T_3 & \sum z_i t_i^2 \end{vmatrix} = 0 \quad (9)$$

Que determina implícitamente el parámetro L , ya que éste está contenido en cada z_i .

Escribiendo por abreviar,

$$\begin{cases} P = T_1 T_3 - T_2^2 \\ Q = -n T_3 + T_1 T_2 \\ R = n T_2 + T_1^2 \end{cases}$$

Entonces, la ecuación (9) se puede escribir en la forma

$$P \sum z_i + Q \sum z_i t_i + R \sum z_i t_i^2 = 0 \quad (10)$$

O bien

$$\sum (P + Q t_i + R t_i^2) \cdot \ln \left(\frac{L}{P_i} - 1 \right) = 0 \quad (11)$$

4. La anterior es una ecuación no-algebraica, que contiene a como incógnita, del tipo

$$f(L) = 0 \quad (12)$$

En donde f es un polinomio no algebraico, que aparece en el lado izquierdo de la ecuación (11). Este tipo de ecuaciones sólo puede resolverse por métodos numéricos, es decir, buscando numéricamente los valores de $f(L)$ correspondientes a distintos valores numéricos, hasta localizar el valor L^* que anula a $f(L)$.

El hecho de que la ecuación $f(L) = 0$ tiene efectivamente una raíz real y única, se desprende de las siguientes consideraciones:

- a. Que $f(L)$ es continua, y está definida desde el valor $L = M = \max(p_1, p_2, \dots, p_n)$ (cn $M > 0$) hasta $L = +\infty$. En las situaciones prácticas que se manejan en demografía, la sucesión p_1, p_2, \dots, p_n suele ser de datos de población en crecimiento. En ese caso $M = P_n$.
- b. Que para $L = M$ se tiene que $f(L) = -\infty$ (porque $M/P_n = 1$)
- c. Que $f(L)$ es monotónicamente creciente con L , o sea que $f'(L) > 0$.
- d. Que para $L \rightarrow +\infty$, se tiene que $f(L) \rightarrow +\infty$.

Hay varios métodos numéricos utilizables para calcular la raíz de la ecuación $f(L) = 0$: el de Newton, el de falsa posición (Regula Falsis), el de bisección y muchos otros. En este caso hemos usado el método de la bisección para resolver esta ecuación (sección 2.2.4).

5. El cómputo numérico de la raíz $L = L^*$ de la ecuación (12) constituye la clave del método que exponemos aquí. Antes de continuar explicando el método, es importante demostrar que esa raíz no depende del origen que se haya escogido para medir

el tiempo. Es decir, que si las fechas censales t_1, t_2, \dots, t_n se dan en la cronología cristiana, o judía, o mahometana, la raíz L^* es una misma. Y, en general, que si se mide el tiempo t desde cualquier momento y en cualesquiera unidades, la ecuación (12) es invariante.

En efecto, al cambiar el origen de tiempo a cualquier otro instante, la nueva fecha τ_i y la original t_i están relacionadas por

$$\tau_i = t_i + \alpha$$

Siendo α el corrimiento que se haya hecho en el origen de tiempo. Entonces:

$$T'_1 = \sum (t_i + \alpha) = T_1 + n\alpha$$

$$T'_2 = \sum (t_i + \alpha)^2 = T_2 + 2\alpha T_1 + n\alpha^2$$

$$T'_3 = \sum (t_i + \alpha)^3 = T_3 + 3\alpha T_2 + 3\alpha^2 T_1 + n\alpha^3$$

$$T'_4 = \sum (t_i + \alpha)^4 = T_4 + 4\alpha T_3 + 6\alpha^2 T_2 + 4\alpha^3 T_1 + n\alpha^4$$

Sustituyendo y haciendo algunas operaciones algebraicas que omitimos, se obtiene

$$P' = T'_1 T'_3 - (T'_2)^2 = P - \alpha^2 T_1^2 - \alpha T_1 T_2 + n\alpha^2 T_2 + n\alpha T_3$$

$$Q' = -nT'_3 - T'_1 T'_2 = Q + 2\alpha T_1 - 2n\alpha T_2$$

$$R' = nT'_2 - (T'_1)^2 = R$$

Luego:

$$P' + Q't_i + Rt_i^2 = P + Qt_i + \alpha.[Q + nT_3 - T_1 T_2] + \alpha.(\alpha + 2t_i).[R - nT_2 + T_1^2]$$

Y los dos corchetes que hay a la derecha son nulos por definición, según las ecuaciones (10). Es decir, que para cualquier cambio de origen en el tiempo, la expresión

$$P + Qt_i + Rt_i^2$$

Es invariante en su forma algebraica y en su valor numérico. Mediante operaciones algebraicas sencillas se muestra que, para un cambio en la escala de tiempo

$$\tau_i = ct_i, \text{ con } c \neq 0,$$

Se obtiene

$$P' + Q'\tau_i + R\tau_i^2 = c^4(P + Qt_i + Rt_i^2)$$

Y por lo tanto, en la ecuación (11) solamente aparece el lado izquierdo multiplicado por el factor c^4 ($c^4 \neq 0$), lo cual, evidentemente, no cambia la raíz L^* de esa ecuación.

6. El parámetro b de la curva logística se calcula recordando que $b = -B$ y que B se obtiene del sistema de ecuaciones normales. Por lo tanto

$$b = -\frac{\begin{vmatrix} n & \sum z_i & T_2 \\ T_1 & \sum z_i t_i & T_3 \\ T_2 & \sum z_i t_i^2 & T_4 \end{vmatrix}}{D}$$

Y así mismo

$$\ln K = \frac{\begin{vmatrix} \sum z_i & T_1 & T_2 \\ \sum z_i t_i & T_2 & T_3 \\ \sum z_i t_i^2 & T_3 & T_4 \end{vmatrix}}{D}$$

También es importante aquí hacer notar que el determinante D es invariante ante cualquier traslación del origen de tiempo:

$$\tau_i = t_i + a,$$

Traslación en la cual

$$D = \begin{vmatrix} n & T'_1 & T'_2 \\ T'_1 & T'_2 & T'_3 \\ T'_2 & T'_3 & T'_4 \end{vmatrix}$$

$$D = n[T'_2 T'_4 - (T'_3)^2] - T'_1(T'_1 T'_4 - T'_2 T'_3) - T'_2[T'_1 T'_3 - (T'_2)^2] \quad (13)$$

Reg. 6474 — 15/12/15 WRP

Sustituyendo según las ecuaciones (13), y después de un largo cálculo algebraico, se encuentra que

$$D' \equiv n[T_2'T_4' - (T_3')^2] - T_1'(T_1'T_4' - T_2'T_3') - T_2'[T_1'T_3' - (T_2')^2]$$

$$D' = n[T_2T_4 - T_3^2] - T_1[T_1T_4 - T_2T_3] - T_2[T_1T_3 - T_2^2]$$

$$\begin{vmatrix} n & T_1 & T_2 \\ T_1 & T_2 & T_3 \\ T_2 & T_3 & T_4 \end{vmatrix} = D$$

Es decir, que el determinante D es invariante en forma algebraica y en valor numérico ante un cambio de origen de tiempo (o cambio de cronología). Se comprueba también que ante un cambio de unidades de tiempo (o cronometría),

$$\tau_i = ct_i$$

El nuevo determinante es $D^1 = c^6 D$

Por lo tanto un cambio en cronología y en cronometría de la forma

$$\tau_i = ct_i + a$$

Modifica el determinante D en la forma

$$D^1 = c^6 D$$

7. Las anteriores observaciones permiten formular un algoritmo explícito, directo y general para calcular los tres parámetros L, K, b de una curva logística

$$p(t) = \frac{L}{1 + Ke^{-bt}}$$

A partir de los datos de población (presumiblemente crecientes), p_1, p_2, \dots, p_n , (siendo $n > 3$), medidos empíricamente (por ejemplo, por recuentos censales), en las fechas t_1, t_2, \dots, t_n y según el criterio de mínimos cuadrados. Para medir y expresar las fechas, se supone que hemos tomado un origen de tiempo (que puede ser cualquiera), y una unidad de medida, de la cual dependerán los parámetros K, b pero no el parámetro L .

8. La tasa relativa de crecimiento poblacional es una variable muy importante en todos los estudios demográficos y está definida por la identidad:

$$r(t) = \frac{1}{p} \cdot \frac{dp}{dt} \quad (14)$$

Que se interpreta como la velocidad del crecimiento de la población, dividida por la población. Al efectuar el proceso de derivación correspondiente, del cual se muestran algunos pasos elementales, se obtiene:

$$\frac{dp}{dt} = \frac{d}{dt} \left(\frac{L}{1 + Ke^{-bt}} \right) = \frac{Lkbe^{-bt}}{(1 + ke^{-bt})^2}$$

Ahora, al reemplazar en la ecuación (14)

$$r(t) = \frac{1}{p} \frac{dp}{dt} = \left(\frac{1 + Ke^{-bt}}{L} \right) \cdot \frac{Lkbe^{-bt}}{(1 + ke^{-bt})^2}$$

Y luego de las correspondientes cancelaciones de términos, la tasa relativa de crecimiento para la curva logística resulta ser:

$$\hat{r}(t) = \frac{K \cdot b}{K + e^{bt}} \quad (15)$$

Siendo t = año que se calcula – año base.

Generalmente el año de base se puede tomar como un año intermedio en la serie de fechas de los datos censales.

III. METODOLOGIA DE LA INVESTIGACION

3.1 Tipo y diseño de investigación

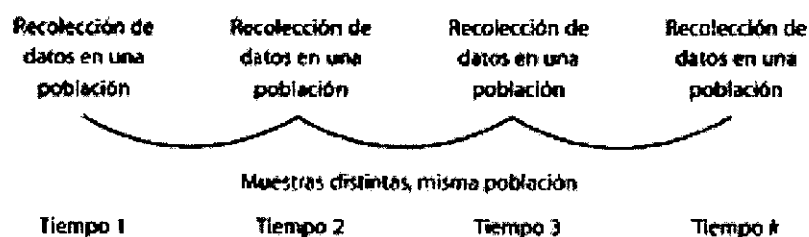
- Tipo de investigación

El tipo de investigación es predictiva

- Diseño de investigación

El diseño del presente trabajo de investigación es no experimental -longitudinal de tendencia. Según Hernández, Fernández & Baptista (2006) el diseño no experimental es una investigación que se realiza sin la manipulación deliberada de variables y en los que solo se observan los fenómenos en su ambiente natural para después analizarlos. Los diseños longitudinales son estudios que recaban datos en diferentes puntos del tiempo para realizar inferencias acerca del cambio, sus causas y sus efectos. De tendencia son los que analizan cambios a través del tiempo, dentro de una población en general.

Cuyo esquema es el siguiente:



3.2 Población y muestra

- Población

En el presente trabajo de investigación se empleara los datos censales de la población peruana, provenientes de los once Censos Nacionales de Población y Vivienda que cubren un periodo de 171 años, desde 1836 hasta 2007.

- Muestra: Comprende la población de los años 1940-2007.

IV. RESULTADOS DE LA INVESTIGACION

4.1. ESTIMACION DE PARAMETROS

A partir de los datos de los censos oficiales que se han hecho desde 1836 hasta 2007, suministrados por el Instituto Nacional de Estadística e Informática INEI (Tabla 1), se calculará la curva logística. Debe considerarse sin embargo, que existe discusión sobre la exactitud de algunas de las cifras correspondientes a los años 1836, 1850, 1862 y 1876 que no se tendrá en cuenta para los cálculos por tener explicables limitaciones de orden técnico, siendo afectados por la crisis de carácter político y económico imperante en dicha época (1836-1850). El Censo de 1862, se ocupó más del Registro Cívico para efectos electorales que del aspecto propiamente estadístico. Las críticas que se hicieron al censo de 1876, estuvieron referidas principalmente a la cédula, dejando de empadronar pueblos y caseríos ejecutándose de forma muy apresurada y sin la debida organización, motivos por los cuales estos primeros años incluían ruido durante el proceso de cálculo. Por lo que el modelo se estimó con los datos a partir del año 1940-2007.

En la Tabla 1 observamos los datos censales de la población peruana obtenidos entre 1940 y el año 2007. Con estos datos se realizarán los diferentes cálculos de la curva logística.

Tabla 1: Datos censales de la población peruana 1940-2007.

Año	Día	Población censada p_i	t_i
1940	9 de junio	7023111	-39.5616438
1961	2 de julio	10420357	-18.5013699
1972	4 de junio	14121564	-7.57808219
1981	12 de julio	17762231	1.5260274
1993	11 de julio	22639443	13.5232877
2005	18 al 20 de julio	27219264	25.5479452
2007	21 de octubre al 4 de noviembre	28220764	27.8410959

Fuente: Instituto Nacional de Estadística e informática (INEI) – Censos Nacionales de Población y Vivienda.

t_i : Es el año censal, contado a partir de 1980: año que se calcula – año base (1980).

Para un mejor entendimiento del procedimiento, se realizarán los cálculos paso a paso, empleando una hoja electrónica de cálculo Excel.

1. Para los siete datos censales de la población peruana (Tabla 1), tómese el origen de tiempo o año base en cualquier momento o fecha y la unidad de tiempo que desee, una fecha conveniente para fijar el origen de tiempo es el valor intermedio entre los años censados 1° de enero de 1980; y unidad de tiempo: año calendario.

En el Cuadro 2 se obtienen los valores de los tiempos elevados a las diferentes potencias.

Cuadro 2:

t_i	t_i^2	t_i^3	t_i^4
-39.5616438	1565.12366	-61918.8649	2449612.08
-18.5013699	342.300687	-6333.03161	117169.7602
-7.57808219	57.4273297	-435.189025	3297.898197
1.5260274	2.32875962	3.553750977	5.423121355
13.5232877	182.879309	2473.129511	33444.84182
25.5479452	652.697504	16675.08007	426014.032
27.8410959	775.12662	21580.37457	600821.2776
2.79726027	3577.88387	-27954.9476	3630365.313

2. Calcúlese los valores numéricos de T_1, T_2, T_3, T_4 :

$$T_1 = \sum t_i, \quad T_2 = \sum t_i^2, \quad T_3 = \sum t_i^3, \quad T_4 = \sum t_i^4$$

Obteniéndose los siguientes resultados mostrados en el Cuadro 3.

Cuadro 3:

T_1	T_2	T_3	T_4
2.79726027	3577.88387	-27954.9476	3630365.313

3. Con estos resultados, calcúlese cada uno de los valores numéricos de P, Q y R donde:

$$P = T_1 T_3 - T_2^2, \quad Q = -n T_3 + T_1 T_2, \quad R = n T_2 - T_1^2$$

Obteniendo lo mostrado en el cuadro 4

Cuadro 4:

P	Q	R
-12879450.3	205692.906	25037.36245

4. Calcúlese cada uno de los valores numéricos de los n coeficientes $P + Qt_i + Rt_i^2$ para $i = 1, 2, \dots, 7$

Cuadro 5:

t_i	$P + Q * t_i + R$
-39.5616438	18169568.66
-18.5013699	-8114744.444
-7.57808219	-13000379.15
1.5260274	-12507251.27
13.5232877	-5518990.381
25.5479452	8717404.798
27.8410959	12254391.78

5. Fórmese la ecuación:

$$\sum (P + Qt_i + Rt_i^2) \cdot \ln\left(\frac{L}{P_i} - 1\right) = 0, \text{ y sea } z_i = \ln\left(\frac{L}{P_i} - 1\right)$$

6. Calcúlese el valor de L en la forma siguiente:

Utilizando el método de bisección (ver sección 3.1), se exploran distintos valores numéricos para L hasta encontrar aquel que hace la $f(L) = 0$

Para seleccionar el primer valor de L_0 y encontrar más rápidamente el valor adecuado evitando un gran número de iteraciones innecesarias, podemos considerar que, como estamos buscando la población límite del país y el último valor de población censado según la Cuadro 1 es $p = 28220764$ habitantes entonces, este valor u otro muy cercano podría servir como punto de partida, en este caso emplearemos $L_0 = 28220765$ habitantes.

Con este valor se realiza una primera iteración para encontrar el resultado de $f(L_0)$ obteniendo un valor negativo, que se muestra en la Cuadro 6.

Cuadro 6:

t_i	$P + Q * t_i + R$	z_i	$P + Q * t_i + R * t_i^2$
-39.56	18169568.66	1.104684233	20071636.02
-18.50	-8114744.444	0.535460082	-4345121.722
-7.58	-13000379.15	-0.001584862	20603.8029
1.53	-12507251.27	-0.529656053	6624541.342
13.52	-5518990.381	-1.400267985	7728065.537
25.55	8717404.798	-3.302425083	-28788576.27
27.84	12254391.78	-17.15556858	-210231058.6
$f_{(L_0)}$			-208919909.9

En una segunda iteración, se selecciona arbitrariamente un valor mayor que L_0 , en este caso usaremos cuatro veces el valor de la última población censada, entonces se tiene que:

$L_1 = 4M = 4 * 28220764 = 112883056$ Habitantes. Con este valor, el resultado para $f_{(L_1)}$ es positivo:

Cuadro 7:

t_i	$P + Q * t_i + R$	z_i	$P + Q * t_i + R * t_i^2$
-39.56	18169568.66	2.712910665	49292416.6
-18.50	-8114744.444	2.285737523	-18548175.87
-7.58	-13000379.15	1.945004781	-25285799.61
1.53	-12507251.27	1.678073577	-20988087.86
13.52	-5518990.381	1.382819174	-7631765.72
25.55	8717404.798	1.146505281	9994550.64
27.84	12254391.78	1.098612289	13462825.4
$f_{(L_1)}$			295963.5837

Iterando varias veces según el método de la bisección (ver ANEXO 1) entre estos dos valores de $L_0 = 28220765$ habitantes y $L_1 = 112883056$ habitantes, se continúa con el valor intermedio $L_2 = 70551910.5$ y así sucesivamente hasta llegar al valor de $L = 96640877.51$ habitantes con el cual se obtienen los resultados de la Cuadro 8:

Puesto que $f(L) = 0.000365468 \approx 0$, puede considerarse que la población límite L para Perú, según los datos censales, es $L = 96640877.51$ habitantes:

Cuadro 8:

t_i	$P + Q * t_i + R$	z_i	$P + Q * t_i + R * t_i^2$
-39.56	18169568.66	2.546347305	46266032.2
-18.50	-8114744.444	2.113146909	-17147647.14
-7.58	-13000379.15	1.765329378	-22949951.25
1.53	-12507251.27	1.490836202	-18646262.98
13.52	-5518990.381	1.184390828	-6536641.585
25.55	8717404.798	0.936273295	8161873.313
27.84	12254391.78	0.885608819	10852597.44
$f(L)$			0.000365468

7. Con estos datos se calculan ahora los valores de:

Cuadro 9:

Σz_i	Σz_i	Σz_i
10.92193274	-86.34351516	6327.692802

8. Luego se forma el determinante:

$$D = \begin{vmatrix} n & T_1 & T_2 \\ T_1 & T_2 & T_3 \\ T_2 & T_3 & T_4 \end{vmatrix} = \begin{vmatrix} 7 & 2.797260274 & 3577.883873 \\ 2.797260274 & 3577.883873 & -27954.94765 \\ 3577.883873 & -27954.94765 & 3630365.313 \end{vmatrix}$$

Al calcular dicho determinante por la regla de Sarrus, con la ayuda de una hoja electrónica de cálculo u otra aplicación de software, se obtiene el resultado:

$$D = 39063460317$$

9. De aquí deducimos

$$b = - \frac{\begin{vmatrix} n & \sum z_i & T_2 \\ T_1 & \sum z_i t_i & T_3 \\ T_2 & \sum z_i t_i^2 & T_4 \end{vmatrix}}{D}$$

$$b = \frac{\begin{vmatrix} 7 & 10.92193274 & 3577.883873 \\ 2.797260274 & -86.34351516 & -27954.94765 \\ 3577.883873 & 6327.692802 & 3630365.313 \end{vmatrix}}{39063460317}$$

$$b = 0.025360343$$

10. De igual forma se procede con el cálculo de k

$$k = \exp \frac{\begin{vmatrix} \sum z_i & T_1 & T_2 \\ \sum z_i t_i & T_2 & T_3 \\ \sum z_i t_i^2 & T_3 & T_4 \end{vmatrix}}{D}$$

$$k = \exp \frac{\begin{vmatrix} 10.92193274 & 2.797260274 & 3577.883873 \\ -86.34351516 & 3577.883873 & -27954.94765 \\ 6327.692802 & -27954.94765 & 3630365.313 \end{vmatrix}}{39063460317}$$

$$k = 4.808620846$$

11. Concluimos que la curva logística que mejor ajusta estos datos censales de Perú está dada por la expresión:

$$\hat{p}(t) = \frac{96640877.505426}{1 + 4.808620846 * e^{-0.025360342*t}} \quad (16)$$

En donde $t = (\text{Año censado}) - 1980$

12. La tasa relativa de crecimiento poblacional para esta curva logística es:

$\hat{r}(t) = \frac{K.b}{K + e^{bt}}$, que para el caso de Perú es igual a:

$$\hat{r}(t) = \frac{0.121948273}{4.808620846 + e^{0.025360342*t}} \quad (17)$$

$t = (\text{Año censado}) - 1980$

De esta manera puede formarse la Tabla 10 que compara los valores históricos de los censos con la estimación retrospectiva que da la curva logística para los diferentes años desde 1940, con las correspondientes tasas relativas instantáneas de crecimiento de la población.

Tabla 10: Tasa relativa de crecimiento y error relativo con datos censales (1940 – 2007)

Fecha censal	t_i	Población censada	Población estimada	Tasa relativa	Error relativo	Cuadrado del error relativo
09/06/1940	-39.56	7023111	6846993	2.36%	-2.51%	0.0006
02/07/1961	-18.5	10420357	9975707	2.24%	-4.27%	0.0018
04/06/1972	-7.578	14121564	12346281	2.16%	-12.57%	0.0158
12/07/1981	1.526	17762231	14584926	2.09%	-17.89%	0.0320
11/07/1993	13.523	22639443	17855016	1.96%	-21.13%	0.0447
20/07/2005	25.548	27219264	21401627	1.81%	-21.37%	0.0457
04/11/2007	27.841	28220764	22097101	1.78%	-21.70%	0.0471

Fuente: Datos Censales suministrados por el INEI

De igual manera puede construirse una tabla con los valores de población proyectada y su respectiva tasa relativa de crecimiento, año por año desde el año 1940 hasta el año 2020 o más, si se desea.

Tabla 11: Población estimada para Perú año por año (1930 – 2060) según la curva logística

Fecha	t_i	Población estimada	Tasa relativa de crecimiento	Fecha	t_i	Población estimada	Tasa relativa de crecimiento
01/01/1930	-50	5342559	2.40%	01/01/1966	-14	12297985	2.21%
01/01/1931	-49	5472011	2.39%	01/01/1967	-13	12572760	2.21%
01/01/1932	-48	5604407	2.39%	01/01/1968	-12	12852740	2.20%
01/01/1933	-47	5739805	2.39%	01/01/1969	-11	13137979	2.19%
01/01/1934	-46	5878263	2.38%	01/01/1970	-10	13428535	2.18%
01/01/1935	-45	6019839	2.38%	01/01/1971	-9	13724460	2.18%
01/01/1936	-44	6164594	2.37%	01/01/1972	-8	14025808	2.17%
01/01/1937	-43	6312587	2.37%	01/01/1973	-7	14332628	2.16%
01/01/1938	-42	6463879	2.37%	01/01/1974	-6	14644970	2.15%
01/01/1939	-41	6618531	2.36%	01/01/1975	-5	14962882	2.14%
01/01/1940	-40	6776606	2.36%	01/01/1976	-4	15286408	2.13%
01/01/1941	-39	6938164	2.35%	01/01/1977	-3	15615592	2.13%
01/01/1942	-38	7103270	2.35%	01/01/1978	-2	15950475	2.12%
01/01/1943	-37	7271987	2.35%	01/01/1979	-1	16291096	2.11%
01/01/1944	-36	7444378	2.34%	01/01/1980	0	16637491	2.10%
01/01/1945	-35	7620507	2.34%	01/01/1981	1	16989694	2.09%
01/01/1946	-34	7800438	2.33%	01/01/1982	2	17347736	2.08%
01/01/1947	-33	7984237	2.33%	01/01/1983	3	17711646	2.07%
01/01/1948	-32	8171968	2.32%	01/01/1984	4	18081448	2.06%
01/01/1949	-31	8363698	2.32%	01/01/1985	5	18457167	2.05%
01/01/1950	-30	8559490	2.31%	01/01/1986	6	18838820	2.04%
01/01/1951	-29	8759411	2.31%	01/01/1987	7	19226424	2.03%
01/01/1952	-28	8963526	2.30%	01/01/1988	8	19619992	2.02%
01/01/1953	-27	9171901	2.30%	01/01/1989	9	20019533	2.01%
01/01/1954	-26	9384601	2.29%	01/01/1990	10	20425053	2.00%
01/01/1955	-25	9601693	2.28%	01/01/1991	11	20836553	1.99%
01/01/1956	-24	9823242	2.28%	01/01/1992	12	21254032	1.98%
01/01/1957	-23	10049312	2.27%	01/01/1993	13	21677483	1.97%
01/01/1958	-22	10279969	2.27%	01/01/1994	14	22106897	1.96%
01/01/1959	-21	10515277	2.26%	01/01/1995	15	22542259	1.94%
01/01/1960	-20	10755300	2.25%	01/01/1996	16	22983552	1.93%
01/01/1961	-19	11000103	2.25%	01/01/1997	17	23430751	1.92%
01/01/1962	-18	11249748	2.24%	01/01/1998	18	23883830	1.91%
01/01/1963	-17	11504297	2.23%	01/01/1999	19	24342757	1.90%
01/01/1964	-16	11763812	2.23%	01/01/2000	20	24807496	1.89%
01/01/1965	-15	12028355	2.22%	01/01/2001	21	25278005	1.87%

Fecha	t_i	Población estimada	Tasa relativa de crecimiento	Fecha	t_i	Población estimada	Tasa relativa de crecimiento
01/01/2002	22	25754239	1.86%	01/01/2032	52	42271871	1.43%
01/01/2003	23	26236146	1.85%	01/01/2033	53	42875908	1.41%
01/01/2004	24	26723671	1.83%	01/01/2034	54	43481674	1.39%
01/01/2005	25	27216754	1.82%	01/01/2035	55	44088980	1.38%
01/01/2006	26	27715327	1.81%	01/01/2036	56	44697636	1.36%
01/01/2007	27	28219322	1.80%	01/01/2037	57	45307451	1.35%
01/01/2008	28	28728661	1.78%	01/01/2038	58	45918230	1.33%
01/01/2009	29	29243265	1.77%	01/01/2039	59	46529780	1.32%
01/01/2010	30	29763046	1.75%	01/01/2040	60	47141905	1.30%
01/01/2011	31	30287915	1.74%	01/01/2041	61	47754408	1.28%
01/01/2012	32	30817774	1.73%	01/01/2042	62	48367094	1.27%
01/01/2013	33	31352523	1.71%	01/01/2043	63	48979765	1.25%
01/01/2014	34	31892055	1.70%	01/01/2044	64	49592223	1.23%
01/01/2015*	35	32436259	1.68%	01/01/2045	65	50204273	1.22%
01/01/2016	36	32985019	1.67%	01/01/2046	66	50815718	1.20%
01/01/2017	37	33538212	1.66%	01/01/2047	67	51426363	1.19%
01/01/2018	38	34095714	1.64%	01/01/2048	68	52036014	1.17%
01/01/2019	39	34657394	1.63%	01/01/2049	69	52644477	1.15%
01/01/2020	40	35223116	1.61%	01/01/2050	70	53251560	1.14%
01/01/2021	41	35792739	1.60%	01/01/2051	71	53857075	1.12%
01/01/2022	42	36366120	1.58%	01/01/2052	72	54460832	1.11%
01/01/2023	43	36943110	1.57%	01/01/2053	73	55062648	1.09%
01/01/2024	44	37523555	1.55%	01/01/2054	74	55662337	1.08%
01/01/2025	45	38107298	1.54%	01/01/2055	75	56259720	1.06%
01/01/2026	46	38694178	1.52%	01/01/2056	76	56854620	1.04%
01/01/2027	47	39284031	1.51%	01/01/2057	77	57446861	1.03%
01/01/2028	48	39876688	1.49%	01/01/2058	78	58036272	1.01%
01/01/2029	49	40471977	1.47%	01/01/2059	79	58622685	1.00%
01/01/2030	50	41069723	1.46%	01/01/2060	80	59205936	0.98%
01/01/2031	51	41669748	1.44%				

*: Fecha actual

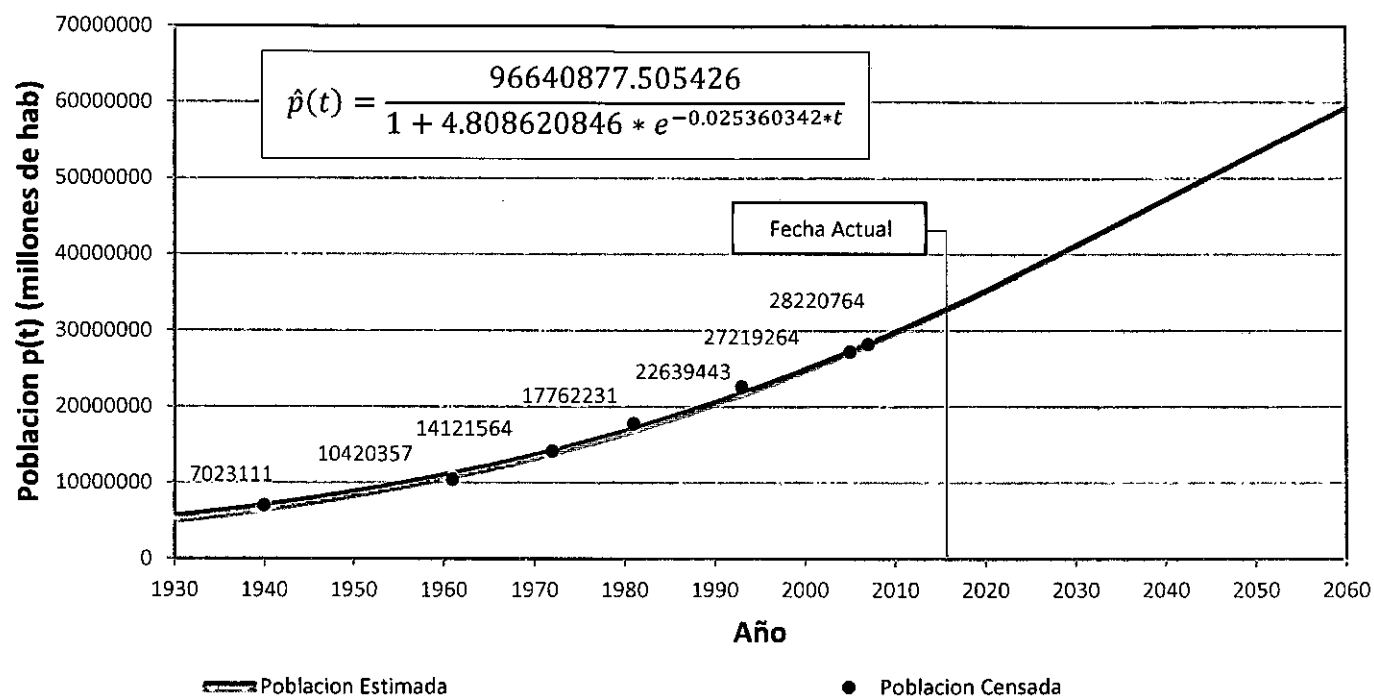
Gráficas

La **gráfica 3** muestra la curva logística calculada con la ecuación (16), desde el año 1930 hasta el año 2060 según los datos censales tomados entre los años 1940 y 2007, estos representados por puntos azules. El año tomado como base para los cálculos fue 1980.

La **gráfica 4** Se observa como la curva toma una forma similar a la mencionada al principio de esta investigación, sobre comportamiento logístico (Gráfica 1, curva intermedia), con crecimiento ascendente desde una población inicial $P(0)$ hasta la población límite L .

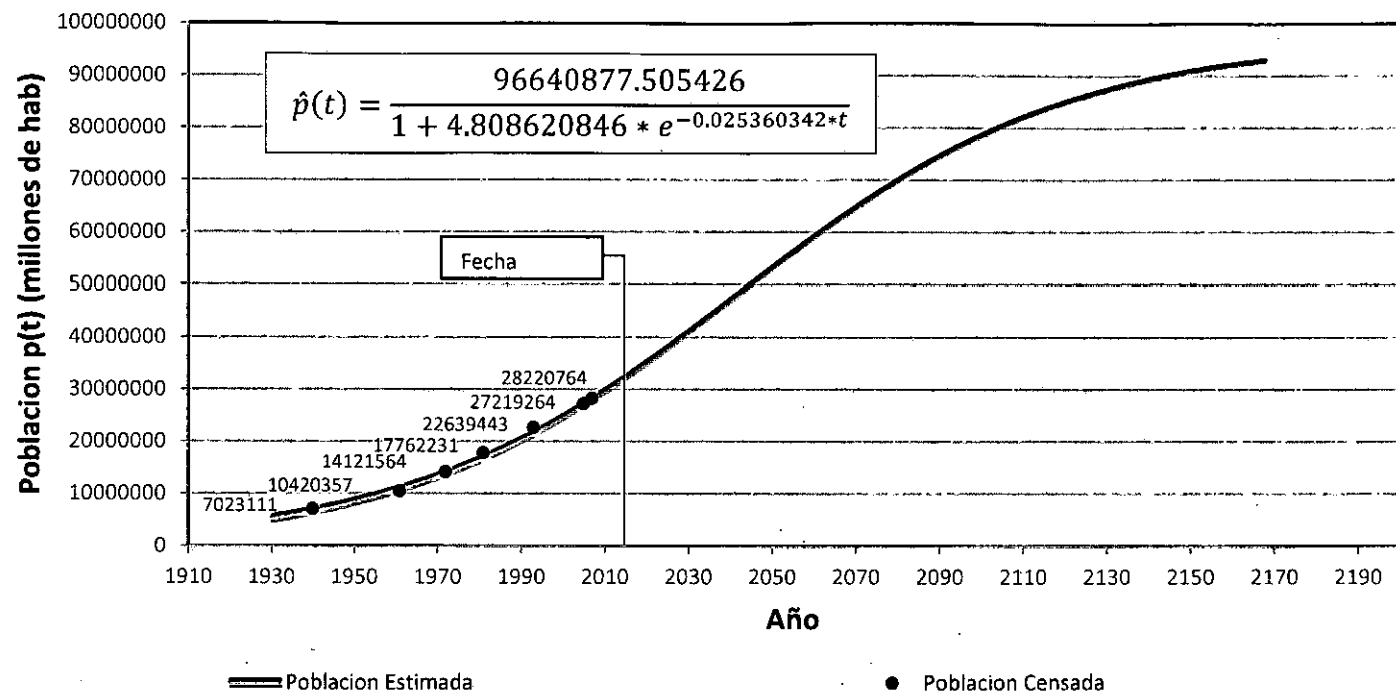
En la **gráfica 5** se observa el comportamiento de la tasa relativa de crecimiento de la población calculada también, a partir de los datos censales tomados entre 1940 y 2007. Se observa como la tasa de crecimiento disminuye desde un valor de 2.36% en el año 1940, pasando por 1.80% en el año 2007 y teniendo un valor más cercano a cero 0.09% en la medida en que la población se acerca a su valor límite hacia el año 2168.

Gráfica 3
Curva Logística de población calculada y proyectada para Perú con censos entre 1940 y 2007
Curva Logística calculada para Perú



Fuente: Datos Censales suministrados por el Instituto Nacional de Estadística e informática (INEI).

Gráfica 4
Curva Logística de población calculada y proyectada para Perú con censos entre 1930 y 2168
Curva Logística a calculada para Perú

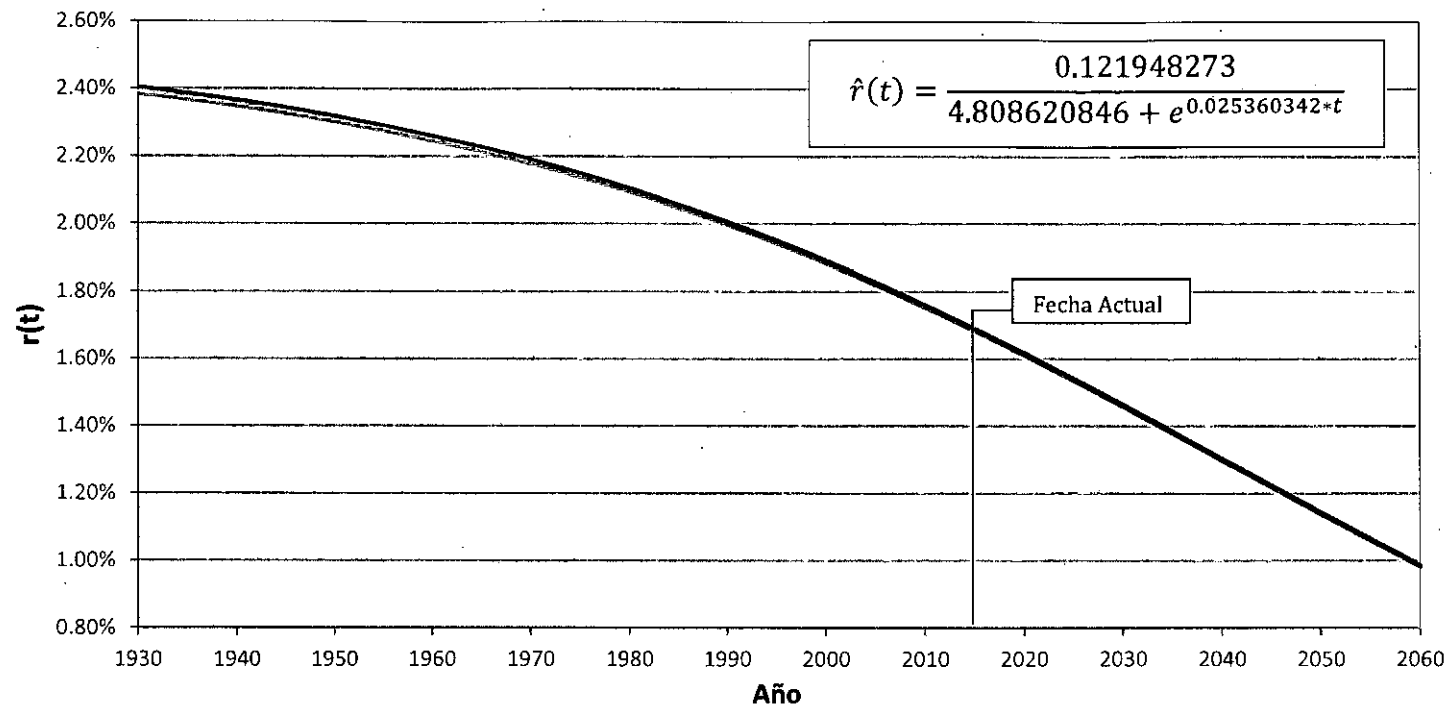


Fuente: Datos Censales suministrados por el Instituto Nacional de Estadística e Informática (INEI).

Gráfica 5

Tasa relativa de crecimiento poblacional para Perú, calculada con los datos censales tomados entre los años 1940 y 2007, proyectadas hasta 2060.

Tasa relativa de crecimiento poblacional $r(t)$ para Perú



Fuente: Datos Censales suministrados por el Instituto Nacional de Estadística e Informática (INEI).

Interpolación entre censos por el método de la parábola de segundo grado, año a año, de 1940 – 2007

Con el fin de determinar los valores de los datos de población intermedios año por año entre los datos censados, se ha realizado con estos, una interpolación por el método de la parábola o polinomial de segundo grado. En el Anexo 3 se describe el método y sus respectivas ecuaciones. En la gráfica 6 se observa su comportamiento y en la Tabla 13 se encuentran los resultados obtenidos.

Gráfica 6
Población estimada de Perú, por el método de interpolación parabólica 1940-2007

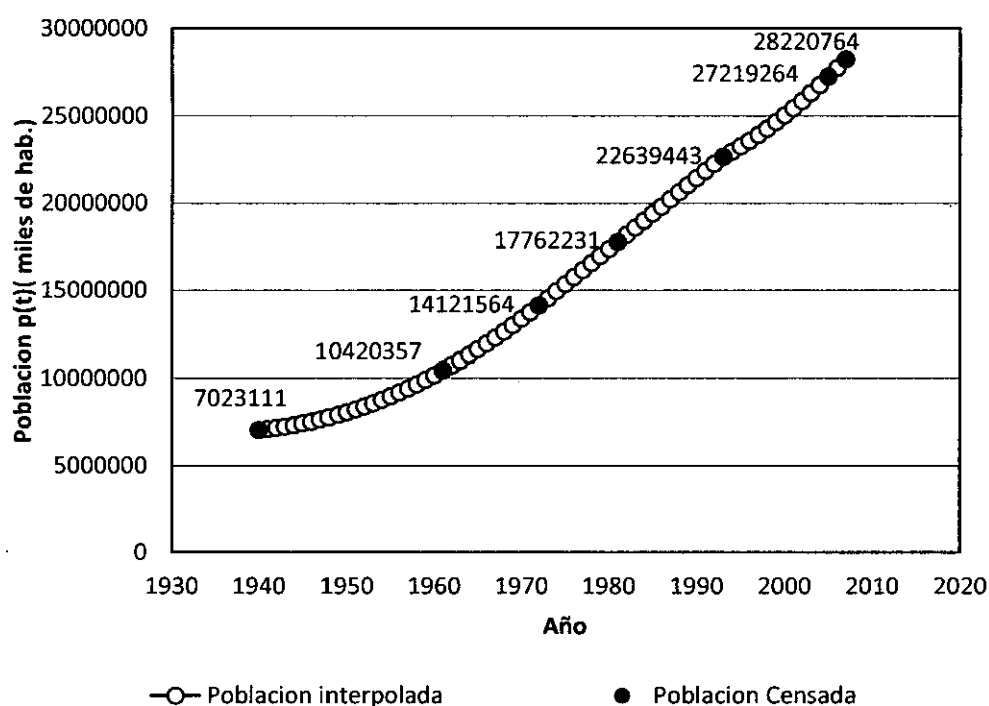


Tabla 13: Población estimada de Perú, por el método de interpolación parabólica

Año	P (t) estimada inter. Parabólica	Población estimada (miles de hab.)	Año	Población Censada (miles de hab.)	P (t) estimada inter. Parabólica
1940	7023111	6776606	1974		14929324
1941	7075697		1975		15333478
1942	7139202		1976		15737814
1943	7213626		1977		16142332
1944	7298969		1978		16547033
1945	7395230		1979		16951917
1946	7502410		1980		17356983
1947	7620508		1981	17762231	17762231
1948	7749526		1982		18167662
1949	7889462		1983		18573275
1950	8040317		1984		18979071
1951	8202090		1985		19385049
1952	8374783		1986		19791210
1953	8558394		1987		20197553
1954	8759237		1988		20604078
1955	8958372		1989		21010786
1956	9174739		1990		21417677
1957	9402026		1991		21824750
1958	9640230		1992		22232005
1959	9889539		1993	22639443	22639443
1960	10149396	11249748	1994		22927518
1961	10420357		1995		23232606
1962	10702237		1996		23554709
1963	10995035		1997		23893825
1964	11298752		1998		24249956
1965	11613388		1999		24623101
1966	11938943		2000		25013260
1967	12275416		2001		25420432
1968	12622808		2002		25844619
1969	12981119		2003		26285820
1970	13350349	14025808	2004		26744035
1971	13730497		2005	27219264	27219264
1972	14121564		2006		27711507
1973	14525353		2007	28220764	28220764

Los valores en **negrita** de la Tabla 13 representan los valores iniciales y finales de los rangos de datos tomados para la interpolación por el método de la parábola.

- **La población promedio (\bar{p})** de Perú durante el período de 68 años comprendido entre 1940 y 2007, calculado con los valores obtenidos por la interpolación parabólica de la Tabla 13 y según la ecuación () fue:

$$\bar{p} = \frac{1}{68} \sum_{t=1}^{68} p(t)$$

$$\bar{p} = 15649283 \text{ habitantes}$$

Con este valor de población promedio se realizan los cálculos siguientes:

- **La varianza poblacional s_p^2** es: 44065863 habitantes, resultado de:

$$s_p^2 = \frac{1}{68} \sum_{t=1}^{68} p(t)$$

- **Error cuadrático promedio** para el rango de 68 años entre 1940 y 2007, es $S_p^2 = 0.001708$ obtenido de la ecuación:

$$S_p^2 = \frac{1}{68} \sum_1^{68} [\varepsilon(t)]^2$$

Con este valor de población promedio se realizan los cálculos siguientes:

Siendo $\varepsilon(t)$ el error relativo entre los datos interpolados de la población, y los valores estimados con la función logística.

En la tabla 14 se muestran los valores correspondientes a los errores relativos y absolutos de las estimaciones de la población para el rango de 68 años comprendidos entre 1940 y 2007.

- El coeficiente de correlación r^2 de la curva logística con los censos realizados es 0.99999999 que se obtiene a partir de la siguiente expresión:

$$r^2 = 1 - \left(\frac{S_p^2}{S_p^2} \right)$$

Donde:

S_p^2 : Varianza poblacional

S_p^2 : Error cuadrático promedio

Tabla 14: Población censada e interpolada vs población estimada

Año	P (t) estimada interp. Parabólica	Población estimada (miles de hab.)	Error relativo $\varepsilon (t)$	Cuadrado del error relativo $\varepsilon (t)^2$	Valor absoluto del Error relativo $ \varepsilon (t) $
1940	7023111	6776606	-4%	0.00123	0.0351
1941	7075697	6938164	-2%	0.00038	0.0194
1942	7139202	7103270	-1%	0.00003	0.0050
1943	7213626	7271987	1%	0.00007	0.0081
1944	7298969	7444378	2%	0.00040	0.0199
1945	7395230	7620507	3%	0.00093	0.0305
1946	7502410	7800438	4%	0.00158	0.0397
1947	7620508	7984237	5%	0.00228	0.0477
1948	7749526	8171968	5%	0.00297	0.0545
1949	7889462	8363698	6%	0.00361	0.0601
1950	8040317	8559490	6%	0.00417	0.0646
1951	8202090	8759411	7%	0.00462	0.0679
1952	8374783	8963526	7%	0.00494	0.0703
1953	8558394	9171901	7%	0.00514	0.0717
1954	8752924	9384601	7%	0.00521	0.0722
1955	8958372	9601693	7%	0.00516	0.0718
1956	9174739	9823242	7%	0.00500	0.0707
1957	9402026	10049312	7%	0.00474	0.0688
1958	9640230	10279969	7%	0.00440	0.0664
1959	9889354	10515277	6%	0.00401	0.0633
1960	10149396	10755300	6%	0.00356	0.0597
1961	10420357	11000103	6%	0.00310	0.0556
1962	10702237	11249748	5%	0.00262	0.0512
1963	10995035	11504297	5%	0.00215	0.0463
1964	11298752	11763812	4%	0.00169	0.0412
1965	11613388	12028355	4%	0.00128	0.0357
1966	11938943	12297985	3%	0.00090	0.0301
1967	12275416	12572760	2%	0.00059	0.0242
1968	12622808	12852740	2%	0.00033	0.0182
1969	12981119	13137979	1%	0.00015	0.0121
1970	13350349	13428535	1%	0.00003	0.0059
1971	13730497	13724460	0%	0.00000	0.0004
1972	14121564	14025808	-1%	0.00005	0.0068
1973	14525353	14332628	-1%	0.00018	0.0133

Fuente: Datos elaborados por el método de interpolación parabólica

Año	P (t) estimada interp. Parabólica	Población estimada (miles de hab.)	Error relativo $\varepsilon(t)$	Cuadrado del error relativo $\varepsilon(t)^2$	Valor absoluto del Error relativo $ \varepsilon(t) $
1974	14929324	14644970	-2%	0.00036	0.0190
1975	15333478	14962882	-2%	0.00058	0.0242
1976	15737814	15286408	-3%	0.00082	0.0287
1977	16142332	15615592	-3%	0.00106	0.0326
1978	16547033	15950475	-4%	0.00130	0.0361
1979	16951917	16291096	-4%	0.00152	0.0390
1980	17356983	16637491	-4%	0.00172	0.0415
1981	17762231	16989694	-4%	0.00189	0.0435
1982	18167662	17347736	-5%	0.00204	0.0451
1983	18573275	17711646	-5%	0.00215	0.0464
1984	18979071	18081448	-5%	0.00224	0.0473
1985	19385049	18457167	-5%	0.00229	0.0479
1986	19791210	18838820	-5%	0.00232	0.0481
1987	20197553	19226424	-5%	0.00231	0.0481
1988	20604078	19619992	-5%	0.00228	0.0478
1989	21010786	20019533	-5%	0.00223	0.0472
1990	21417677	20425053	-5%	0.00215	0.0463
1991	21824750	20836553	-5%	0.00205	0.0453
1992	22232005	21254032	-4%	0.00194	0.0440
1993	22639443	21677483	-4%	0.00181	0.0425
1994	22927518	22106897	-4%	0.00128	0.0358
1995	23232606	22542259	-3%	0.00088	0.0297
1996	23554709	22983552	-2%	0.00059	0.0242
1997	23893825	23430751	-2%	0.00038	0.0194
1998	24249956	23883830	-2%	0.00023	0.0151
1999	24623101	24342757	-1%	0.00013	0.0114
2000	25013260	24807496	-1%	0.00007	0.0082
2001	25420432	25278005	-1%	0.00003	0.0056
2002	25844620	25754239	0%	0.00001	0.0035
2003	26285820	26236146	0%	0.00000	0.0019
2004	26744035	26723671	0%	0.00000	0.0008
2005	27219264	27216754	0%	0.00000	0.0001
2006	27711507	27715327	0%	0.00000	0.0001
2007	28220764	28219322	0%	0.00000	0.0001

Fuente: Datos elaborados por el método de interpolación parabólica

V. CONCLUSIONES

5.1. CONCLUSIONES

Con base en los datos censales suministrados por el Instituto Nacional de Estadística e Informática (INEI) y los cálculos realizados se pueden afirmar las siguientes conclusiones sobre la demografía de Perú:

- La curva logística de la población peruana que mejor ajusta los datos censales de Perú es:

$$\hat{p}(t) = \frac{96640877.505426}{1 + 4.808620846 * e^{-0.025360342*t}}$$

Siendo los valores de los parámetros: $k = 4.8086$, $b = 0.0254$ y $L = 96640877$ habitantes.

- Luego de realizar el cálculo para deducir la ecuación de la curva con la cual se comparan los datos censales disponibles a la fecha, de Perú, se observa como la curva de la Gráfica 4 toma una forma sigmoideal similar a la curva intermedia de la Gráfica 2, sobre comportamiento logístico con crecimiento ascendente desde una población inicial $P_{(0)}$ hasta la población límite L .
- En la gráfica 3 se presentan las poblaciones calculadas y proyectadas. Es posible observar que los resultados proyectados que derivan de la función de la curva logística se asemejan a los calculados, lo que significa que el modelo reproduce de manera adecuada la dinámica demográfica del Perú, en el periodo 1940-2007.
- Con este el método original la curva logística se ajusta muy bien a los datos de los censos realizados en Perú. Y provee un modelo aceptable para explicar su crecimiento poblacional.
- El coeficiente de correlación $r^2=0,99999999$ resultante de la comparación entre los valores censados (incluidos los interpolados) durante todo el rango de tiempo entre 1940 y 2007, con la curva logística, es altamente satisfactorio, lo que muestra un excelente ajuste.
- A mediados del siglo XXII la población estará en un valor cercano a 90'783,900 habitantes y estará acercándose asintóticamente al límite máximo de saturación de habitantes hacia el 2168 (valor calculado por fuera de las gráficas mostradas), es decir cerca de los 96'000,000 habitantes. A partir del valor de la población límite (96'640878 habitantes), cifra que la historia censal permite pronosticar, a la

fecha, una de las interpretaciones que puede tener este valor es la indicación de que, a partir de este número de habitantes, se iniciará un proceso de hacinamiento, lo cual repercutirá en el incremento progresivo de problemas sociales por cada habitante nuevo en el País.

- La tasa de crecimiento relativo poblacional ha venido reduciéndose desde 2,36% en el año 1940 a 1,80% en el año 2007. Para el año 2060 se espera un valor de 0,98%; lo cual indica que, de continuar la tendencia, el crecimiento se estabilizará al llegar al límite de población.
- En el 2015, la población peruana es de 32 millones de personas y hacia el año 2060, la población habrá alcanzado los 59 millones de personas, según los datos existentes a la fecha. La población seguirá aumentando durante los próximos años, aun con tasas porcentuales declinantes.
- Debido a que los valores censados se encuentran distribuidos a lo largo de una curva, la interpolación parabólica parece ser la más adecuada, además de ser un método factible de aplicar. El método usado aquí para interpolar puede afinarse o modificarse en varios sentidos, con cambios solo insignificantes en los resultados de esta investigación. Es decir, este es un método estadísticamente muy robusto.
- El método de bisección presenta varios puntos a favor, siendo el más importante de estos, su sencillez conceptual, su facilidad de interpretación e implementación, proporcionando una estimación predeterminada de la precisión de la solución calculada.
- Los resultados de esta investigación, pueden servir de insumo para las instituciones privadas, estatales y gubernamentales para su proceso de planeación estratégica de los recursos y políticas de nuestro Perú.

5.2. RECOMENDACIONES

- Por lo anteriormente expuesto, la recomendación es utilizar la curva logística para realizar predicciones poblacionales, ya sea a corto, mediano y largo plazo. Lo cual ayudara al estado de un País a estar preparado para cualquier escenario, debido a la preocupación de que los recursos naturales y materiales del mundo que no puedan ser capaces de continuar proporcionando las necesidades básicas de las personas.

VI. BIBLIOGRAFIA

Hernández, R., Fernández, C. & Baptista, P. (2006). Metodología de la Investigación (4ª Edición). México: McGraw-Hill.

Informe del Fondo Alemán de la Población de la Tierra (2014). Disponible en http://sp.ria.ru/spanish_ruvr_ru/news/2014_01_01/La-poblaci-n-mundial-alcanza-7-2-mil-millones-de-personas-0330/

Instituto Nacional de Estadística e Informática. Historia de los censos en el Perú, 2007. Disponible en http://censos.inei.gob.pe/Censos2007/Documentos/Historia_Censos.pdf

Veres, E. (1955). Nueva determinación de la curva logística de la población de España". Disponible en http://www.ine.es/ss/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3B+filename%3D1007%2F262%2F108_1.pdf&blobkey=urldata&blobtable=MungoBlobs&blobwhere=1007%2F262%2F108_1.pdf&ssbinary=true.

Medina, J. & Suárez, J. (2000). Modelo logístico adaptado a la población mexicana en el periodo 1900-1990. *Investigación y ciencia de la Universidad Autónoma de Aguascalientes*, 16-23. Disponible en <http://www.uaa.mx/investigacion/revista/archivo/revista22/Articulo%204.pdf>

- Ordorica, M. (2008). Evolución de la población de México, 1980-2005, conforme a la hipótesis de una tasa de crecimiento demográfico logística. *Estudios demográficos urbanos*, 23(3), 445-479. Disponible en <http://www.redalyc.org/articulo.oa?id=31211538001>
- Poveda, G. & Manrique, J. (2007). Aplicación de la curva logística a los censos de la ciudad de Medellín. *Ecos de Economía*, 11(25), 7-60. Disponible en <http://publicaciones.eafit.edu.co/index.php/ecos-economia/article/view/1937/1948>

VII. ANEXOS

ANEXO 1. Método de bisección.

Este es uno de los métodos más sencillos para resolver ecuaciones en una variable. Se basa en el Teorema de los Valores Intermedios, el cual establece que toda función continua f en un intervalo cerrado $[a, b]$ toma todos los valores que se hallan entre $f(a)$ y $f(b)$. Esto es, que todo valor entre $f(a)$ y $f(b)$ es la imagen de al menos un valor en el intervalo $[a, b]$. En caso de que $f(a)$ y $f(b)$ tengan signos opuestos $f(a) \cdot f(b) < 0$, el valor cero sería un valor intermedio entre $f(a)$ y $f(b)$, por lo que con certeza existe un punto p en $[a, b]$ que cumple $f(p) = 0$. De esta forma, se asegura la existencia de al menos una solución de la ecuación

$$f(x) = 0.$$

El método consiste en lo siguiente:

Verificar la continuidad de la función $f(x)$ en el intervalo $[a, b]$ - ver numeral 4 del Ajuste de la logística por mínimos cuadrados.

Luego se verifica que $f(a) \cdot f(b) < 0$

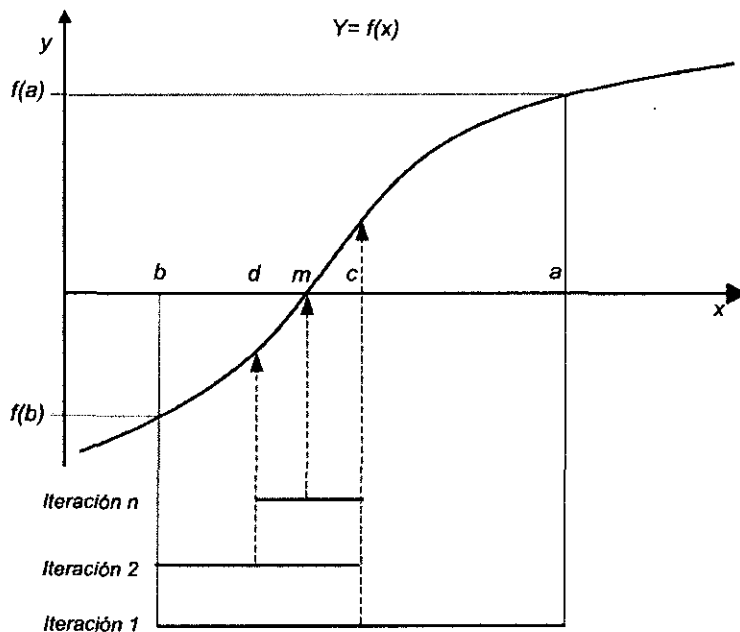
$$L = \lim_{n \rightarrow \infty} L_n \qquad L_{n+1} = \frac{1}{2}(L_n + L_{n+1})$$

$$L_n L_{n-1} < 0$$

Se calcula el punto medio c del intervalo $[a, b]$. A continuación se calcula $f(c)$. En caso de que $f(c)$ sea igual a cero, la raíz buscada será c . En caso de que no lo sea, se verifica si $f(c)$ tiene signo opuesto con $f(a)$ o con $f(b)$. Se redefine el intervalo $[a, b]$ como $[a, c]$ o $[c, b]$ según se haya determinado en cuál de estos intervalos ocurre un cambio de signo. Con este nuevo intervalo se continúa sucesivamente encontrando el centro y encerrando la solución en un intervalo cada vez más pequeño, hasta alcanzar la precisión deseada. En la Gráfica 3, se ilustra el procedimiento descrito en el cual, el cambio de signo ocurre inicialmente entre $[a, b]$; se busca el punto intermedio c y se redefine el rango $[a, c]$. Se verifica el cambio de signo y finalmente, entre $[d, c]$ se encuentra el punto intermedio m como raíz de $f(x)$.

Gráfica 3

Ilustración del método de la bisección



Desde la primera iteración del algoritmo de bisección, se observa que la raíz m se halla a una distancia menor o igual que $\frac{b-a}{2}$. En la segunda iteración, el nuevo intervalo $[a, c]$ mide $\frac{b-a}{2}$ y de nuevo la distancia entre el nuevo punto medio y m es menor o igual que $[d, c] \left(\frac{b-a}{2} \right) \div 2 = \frac{b-a}{4}$. Si se continúa así

sucesivamente, en la n -ésima iteración, al aproximar la raíz p de la función $f(x)$, con el punto de acercamiento m (renombrado como el punto medio durante n iteraciones m_n) se tiene que:

$$|m_n - p| \leq \frac{b-a}{2^n} \quad (16)$$

Y según esta desigualdad,

$$n = \frac{\ln\left(\frac{b-a}{\varepsilon}\right)}{\ln 2} \quad (17)$$

n en su valor entero, es el número de iteraciones necesarias para aproximar la solución de la ecuación mediante el punto medio con un parámetro máximo ε de error.

ANEXO 2. Hoja de cálculo para determinar la ecuación de la Curva Logística

CURVA LOGÍSTICA PARA LOS CENSOS DE PERÚ ENTRE 1940 Y 2007

Año	Población censada p_i	t_i	t_i^2	t_i^3	t_i^4	$P + Q + t_i + R + t_i^2$	$z_i = \ln((L/p_i) - 1)$	$P + Q + t_i + R + t_i^2 + z_i$	$z_i + t_i$	$z_i + t_i^2$
1940	7023111	-39.561644	1565.1237	-61918.865	2449612.08	18169568.66	2.546347305	46266032.2	-100.73769	3985.3484
1961	10420357	-18.50137	342.30069	-6333.0316	117169.76	-8114744.444	2.113146909	-17147647.14	-39.096113	723.33164
1972	14121564	-7.5780822	57.42733	-435.18902	3297.8982	-13000379.15	1.765329378	-22949951.25	-13.377811	101.37815
1981	17762231	1.5260274	2.3287596	3.55375098	5.42312135	-12507251.27	1.490836202	-18646262.98	2.2750569	3.4717991
1993	22639443	13.523288	182.87931	2473.12951	33444.8418	-5518990.381	1.184390828	-6536641.585	16.016858	216.60058
2005	27219264	25.547945	652.6975	16675.0801	426014.032	8717404.798	0.936273295	8161873.313	23.919859	611.10324
2007	28220764	27.841096	775.12662	21580.3746	600821.278	12254391.78	0.885608819	10852597.44	24.65632	686.45897
		2.7972603	3577.8839	-27954.948	3630365.31		10.92193274	0.000365468	-86.343515	6327.6928
		T_1	T_2	T_3	T_4		Σz_i		$\Sigma z_i + t_i$	$\Sigma z_i + t_i^2$

n	7
-----	---

P	-12879450
Q	205692.906
R	25037.3624

L	$f(L)$
96640877.51	0.000365468

Para calcular D		
7	2.797260274	3577.8839
2.79726027	3577.883873	-27954.948
3577.88387	-27954.94765	3630365.3

Para calcular b		
7	10.9219327	3577.883873
2.79726027	-86.343515	-27954.94765
3577.88387	6327.6928	3630365.313

Para calcular k		
10.92193274	2.7972603	3577.8839
-86.34351516	3577.8839	-27954.948
6327.692802	-27954.948	3630365.3

D	39063460317
-----	-------------

b	0.02536034
-----	------------

k	4.8086208
-----	-----------

Ecuación de la curva logística

$$\hat{p}(t) = \frac{L}{1 + k * e^{-b*t}}$$

$$\hat{r}(t) = \frac{k * b}{k + e^{b*t}}$$

Ecuación de la curva logística para la población de Perú

$$\hat{p}(t) = \frac{96640877.505426}{1 + 4.808620846 * e^{-0.025360342*t}}$$

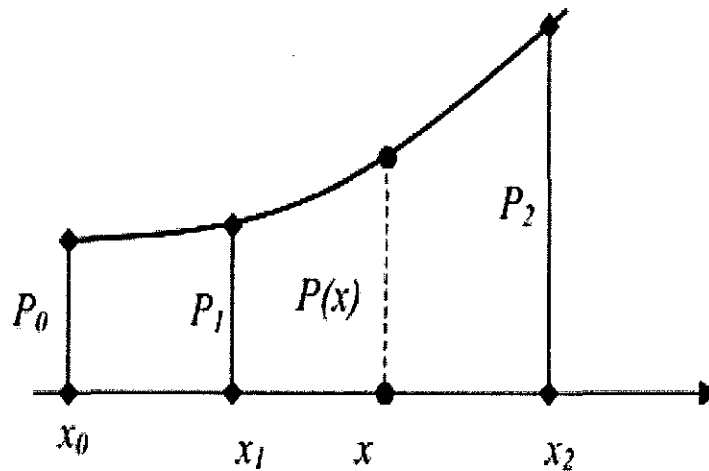
$$\hat{r}(t) = \frac{0.121948273}{4.808620846 + e^{0.025360342*t}}$$

ANEXO 3. Interpolación por el método de la parábola.

Cuando se tienen tres puntos en un plano, puede determinarse la ecuación de un polinomio de segundo grado (una parábola) que contenga dichos puntos. De esta forma, pueden encontrarse los puntos intermedios que pertenezcan a la curva, entre los puntos extremos conocidos. Si se tienen cuatro puntos conocidos, el polinomio será de tercer grado y así sucesivamente. Este procedimiento se denomina interpolación polinomial de Lagrange, y específicamente para tres puntos conocidos, interpolación cuadrática o parabólica.

Gráfica 7

Interpolación parabólica



El teorema que define el n -ésimo polinomio interpolante de Lagrange es el siguiente [12]:

Si x_0, x_1, \dots, x_n son $n+1$ números distintos y f es una función cuyos valores están dados en esos números, entonces existe un único polinomio P de grado a lo más n , con la propiedad de que

$$f(x_k) = P(x_k) \quad \text{Para cada } k=0, 1, \dots, n.$$

Este polinomio está dado por

$$P(x) = f(x_0) L_{n,0}(x) + \dots + f(x_n) L_{n,n}(x) = \sum_{k=0}^n f(x_k) L_{n,k}(x) \quad (19)$$

Dónde:

$$L_{n,k}(x) = \frac{(x-x_0)(x-x_1) \dots (x-x_{k-1})(x-x_{k+1}) \dots (x-x_n)}{(x_k-x_0)(x_k-x_1) \dots (x_k-x_{k-1})(x_k-x_{k+1}) \dots (x_k-x_n)} \\ = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{(x-x_i)}{(x_k-x_i)} \quad (20)$$

Para cada $k=0,1, \dots, n$

$L_{n,k}(x)$ Suele escribirse como $L_k(x)$.

Al aplicar esta ecuación a los puntos de la gráfica se obtiene:

$$p(x) = \frac{(x-x_0)(x-x_1)}{(x_0-x_1)(x_0-x_2)} p_0 + \frac{(x-x_0)(x-x_1)}{(x_1-x_0)(x_1-x_2)} p_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} p_2$$

Al evaluar el polinomio en los puntos conocidos:

$$P(x_0) = p_0$$

$$P(x_1) = p_1$$

$$P(x_2) = p_2$$

Otra forma de escribir el polinomio de interpolación es, mediante los determinantes de las ecuaciones simultáneas de segundo grado que resultan al evaluar cada uno de los tres puntos conocidos de la parábola de la Gráfica 8. Para el caso del punto p, la ecuación es:

$$P(x_\Omega) = a_\Omega + a_1 x_\Omega + a_2 x_\Omega^2$$

Y la ecuación para cualquier punto $p(x)$:

Donde

$$D = \begin{vmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{vmatrix}, D_0 = \begin{vmatrix} p_0 & x_0 & x_0^2 \\ p_1 & x_1 & x_1^2 \\ p_2 & x_2 & x_2^2 \end{vmatrix}, D_1 = \begin{vmatrix} p_0 & 1 & x_0^2 \\ p_1 & 1 & x_1^2 \\ p_2 & 1 & x_2^2 \end{vmatrix}, D_2 = \begin{vmatrix} p_0 & 1 & x_0 \\ p_1 & 1 & x_1 \\ p_2 & 1 & x_2 \end{vmatrix} \quad (21)$$

Los cuales, luego de ser solucionados por la regla de Sarrus, se traducen en las siguientes expresiones, que al reemplazarse en la ecuación (21), permiten realizar la interpolación correspondiente. Este procedimiento puede hacerse con una calculadora programable o una hoja de cálculo:

$$D = x_{\Omega}x_1 (x_1 - x_{\Omega}) + x_1x_2 (x_2 - x_1) - x_2x_{\Omega} (x_2 - x_{\Omega}) \quad (22)$$

$$D_{\Omega} = p_{\Omega}x_1x_2 (x_2 - x_1) - p_{\Omega}x_2x_{\Omega} (x_2 - x_{\Omega}) + p_2x_{\Omega}x_1 (x_1 - x_{\Omega}) \quad (23)$$

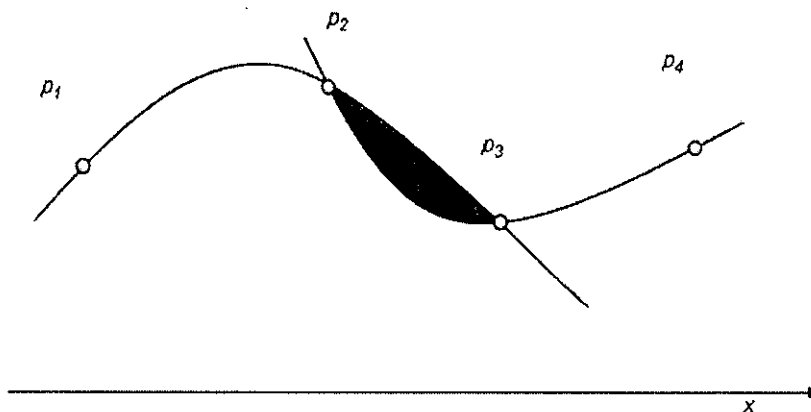
$$D_1 = p_{\Omega} (x_2^2 - x_1^2) - p_1 (x_2^2 - x_{\Omega}^2) + p_2 (x_1^2 - x_{\Omega}^2) \quad (24)$$

$$D_2 = p_{\Omega} (x_2 - x_1) - p_1 (x_2 - x_{\Omega}) + p_2 (x_1 - x_{\Omega}) \quad (25)$$

Existen dos formas de realizar la interpolación parabólica con tres puntos conocidos: una de ellas es, tomando el tercer punto como origen del rango siguiente y así sucesivamente (ver Gráfica 9). La otra consiste en tomar el punto intermedio como origen del rango siguiente, lo cual se muestra en la Gráfica 8, donde P_1 , P_2 y P_3 son los puntos conocidos del primer rango de interpolación. El segundo rango lo conforman los puntos P_2 , P_3 y P_4 y así sucesivamente pueden observarse como los valores de la función entre P_2 y P_3 pueden ser diferentes según la parábola seleccionada y se encuentran en el área sombreada. Por lo tanto, los valores obtenidos por este método, presentan alguna discrepancia en este trayecto

Gráfica 8

Interpolación parabólica con un punto traslapado



En ambos casos, los valores resultantes de la interpolación, son aproximados y los métodos se consideran válidos.

El método a utilizar en este ejercicio será el de interpolación mediante parábolas consecutivas representada en la Gráfica 9:

Gráfica 9

Interpolación parabólica mediante parábolas consecutivas

